

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCGAAGACTAACATTTTGTGAAGTTGTAAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGACTTGGAATACTGAGTTGTTTAGGACTTTCATTGTGGCAAACTT
CCAGAAAACAACCTTTTTGCTGCACATGTAAGTGGAGCTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTGTTTCAGACCATCCTTTCCTACCAAAATGCAGCCCAAAATCCATGGCAAAACAGTC
TTCGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTGGGACTGATTAGAACAGAACTCCATTGGAAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAAATGGTCTATGCTATTTCCCTCTTT
GGTTTTTCCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTACGGGTGGAAGCCAATTTACA
TGGATTAAACCTCTATGACACTGCACCTTGCCCTATTAAACAATGAACGAACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAGG
TTCACAGAAGTTGCTTATTCCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWFFQQGLSFLPSALVINTSAAFI FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERT
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

[illegible]

APP ID=10063713

FIGURE 4

MSFLIDSSIMITSQILFFGFWLFFMRQLFDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRVYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLSCLLWLTIFYFFWKLGP
FPILSPKHGILSIEQLISRVGVIGVITMALLSGFVAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWMIKSVTTSASGSENLTLIQQEVDLEELSRQ
LPLETADLYATKERIEYSKTFKGYFNFLGYFFSIYCVWKIPMATINIVFDRVGKTDPFVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKEFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSAALSSILFLYLAHQAEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCGCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCACTCGCCTCTCATATCCAGCTGGCCATCTGAGGT
 GTTTCCTCGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCTCGTCCAGGCTCTTTGGCTGCAGAGAGCTTCCATCCAGGT
 GTCATCGAGAATTATGGGGATCACCTTGTGAGCAAAAGGCGAACAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAGCTTCAGCTATGGCTGGGTGGGATGGATTCTGTGCTCATCTCTAG
 GATTAGCCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTCTCTGATTGGAGGTTCCAGTGA
 GCCGACAGTTTGACGCTATTGTTACAACATCATCTGATACTTGGACTAACTCGTCGATCCAGAA
 ATTATCACCCACCAAGATCCCATATTCAACACTCAAACTGCAACACAAACACAGAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCTACTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTGTGTACAGAAAGTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAAAAGCAGCAATCAAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCACGCGCTCTGCTAGTGTCTCTCTCTCTCTTTTGGTGTG
 CAGCTGGTCTTTGGATTTTGCTATGTCAAAGGTATGTGAAGGCCCTTCCCTTTACAAACAGAAT
 CAGCAGAGGAATGATCGAAACCAAGTAGTAAAGGAGGAGAGGCCAATGATAGCAACCTAA
 TGAGGAATCAAGAAAACTGATAAAAACCCAGAAGAGTCCAGAGTCCAAGCAAACTACCCTGC
 GATGCCCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTCTTT
 CATGCTCCTTACCCTGCCCGAGCTGGGGAATCAAAGGGGCAAGAGAACCAAGAGAAAGTCCA
 CCCTTGGTTCTTAAGTGAATCAGCTCAGGACTGCCATTGGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACTCTGTCTGATCCTTCTCTACTCTCCAAAGCTTCCCGAGCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCACTGGGAGAAAGGATTTTGCAGAGTCAAGGAC
 CTAAACATCTCATCAGTATCCAGTGTAAAGGGCTCCTGGCTGTCTGAGGCTAGTGGGTG
 AAAGCCAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCGCCAGCTCAGACCTTCTCTTA
 GCTCTGAAAGAGAAACAGTATCCCACTGACATGTCTTCTGAGCCGGTAGAGCAAAAGAAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCAATGGAGATTCTCATACTTGAGACTTAATCTCTGTAA
 GCTAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTGAGCAGGACTGTAAAC
 ACAGACAGGCTCAAAGTGTTTCTCTGAACACATTGAGTTGGAACTCACTGTTTGAACACACACA
 CTTACTTTTCTGGTCTACCACTGCTGATATTCTCTAGGAATATACTTTTACAGATTAACA
 AAAATAAAAACTCTATAAATTCTATTTTATCTGAGTTACAGAAATGATTACTAGAGAAAT
 ACTCAGTAATTTGTTAAAAAGTAATAAATCAACAACATTTGCTGAATAGCTACTATATGTC
 AAGTGTGTGCAAGGATTAACACTCTGAATTGAATATTATTCCTCAAAAAATGCAATAGTAG
 AAGCTATCTGGGAGCTATTTTTCTAGTTTGTATTTCTTAGCTATCTACTTCCAACTAAT
 TTTTATTTTGTCTGAGACTAATCTTATTCATTTTCTCTAATATGCAACCATATTAACCTTAAT
 TTTTATTAACATACCTAAGAAGTACATTTTACCTCTATATCCCAAGGACATTTTAAAGTGC
 ATTATCAAAATGTACTAGGCTCCTTTTCCACAAGAGGGACTGAGAGATGACAGAAATAT
 TGTGACAAAAAATTAAAGCATTAGAAAATCT

FIGURE 6

MARCFSLVLLLTSIWTRLLVQGSRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTTPPAPASTSIPR
RKKLICVTEVEMETSTMSTETEPFVENKAAFKNEAAGPGGVPTALLVLALLFFGAAGLGFCYVK
RYVKAPFPPTNKNQKEMIETKVVKEEKANDSNFNEESKKTDRNPEESKSPSKTTVRCLAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCGCCACCGCGCGCTCCGCACTGACCCCGAGCCGCGG
GGCTCCCTCGGGCGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCACTCGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTGGGGCCACCCTGCTGTGCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGCGCGG
GCTCTCAGCTACCCGAGGAGGAGGCCAACCCTCAATGAGATGTTCCGCGAGGTTGAGGAATGAT
GGAGGACACGACACAAATTGCGCAGCGGTTGGAAGATGGAGGACAGAAGCTGCTGCTA
AAGCATCATCAGAACTGAACCTGGCAAACTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTGGAAATTAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACACAGAC
TGGCAAAATGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGCGAGAAGGAGCC
ACGAGTGCAATCATCGACGAGGACTGTGGGCCACGATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCGGGCGAGCAATGGGACCATCTGTGACA
ACGAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCTTCCAGAGAGGCTGCTGTTCCTGTGTGC
ACACCCCTGCCGTGGAGGGCGAGCTTTGCCATGACCCCGCAGCCGGCTTCTGGACCTCATCAC
CTGGAGACTAGAGCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTGTGGGGAGCCGTGACCAAGATGGG
GAGATCTCTGTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGAGGAGGTTGG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTCGCGCTG
CCGCGCGTGCATGCTGGGAGGGGAAGAGATTAGATCTGGACCAAGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTATTTCCCAAGTGTGTGCTTTAGGCGTGGGCTGACCAAGCTTCTTCCTA
CATCTTCTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGCT
CCCCAGGCTGTCTCCAGGCTTCACAGTCTGGTGTGGGAGAGTCAAGCAGGTTAAACTGCA
GGAGCAGTTTGCACCCCTGTCCAGATTAATGGCTGCTTGGCTTACAGTGTGCGACAGCGCG
TTTGTCTACATGGCTTTGATAATTGTTGAGGGAGGAGATGGAACAATGTGAGTCTCCCTC
TGATTGGTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAACCTGGCAAAATG
CAACAATGAATTTTCCAGCAGTCTTTCATGGCATAGGTAAGCTGTGCTTCAGCTGTGTG
AGATGAATGTTCTGTTCAACCCTGCATTACATGTGTTTATTATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCGAGGCGAGCATTTTCATATCCAGATCAATCCCTCTCTCAGCACAGCTGGGG
AGGGGGTCAATTGTTCTCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGGC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCACACAGGCTTGGTGCCACCAAAGTGTCTCCCAAAGGAAGGAGATGGGAT
TTTTCTGAGGCATGCATCTCGGAATTAAGTCAAACTAATTTCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTTCTCACAGTGTGGGCGACCGCTCCTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAGGTATATGACTGAGCGTAGCA
TACAGTTAACTGCGAAGACAGTACTTAGGTAAATTGTAGGCGAGGATTATTAATGAATTTGC
AAATCACTTAGCAGCAACTGAAGACAATTATCAACCACTGGAGAAATCAACCGAGCAGCGG
GTGTGAACAATGGTTGAATATGCACTGCGCACTGAACTCTACGCCACTCCACAATGATG
TTTTAGGTTGATGGACTGTTGCCACCATGTATTCAACCAAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTAGTTTAAATATGTATAAACATAAGTTGCAATTTAGAA
ATCAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPATATSA PVKPGPALSYFQEEATLWEMFREVEELMEDTQHKL
RSAVEEMEAEAAAAKASSEVNLANLPSPYHNETNTDTKVGNNTHVHREIHKITNNQTCGMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQMLCTRDSGCCDQLCVWGHC
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSPMEEVROELEDLE
RSLTEEMALGEPAAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGAGCGTGGGGGCTGTGAGAAAGTGCAATAAATACATCATGCAACCCAC
GGCCACACTTGTGAACCTCTGTCGCCAGGGCTGATGTGGCTCTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAAATCTATGGGGTCTGGGGCTCTTCTGGACCCCT
AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTGCTCTCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCCGGTCACTTGG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCGTAGCCCCGTGCATCATGTGCTGTTT
CAAGTCTGCCTCTGGTGTCTGGAATAATTTATCAAGTCTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCACAGCCAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTGAGGGTGGTGGTCTGGACAAAGTCACAGACCTGTGCTGTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGGCTCTGTCTCTTTTTTTCTCCGGTGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT
GTCTACGCCACGGCTCTTTCAGCGTTTCGGCATGTGTGTGACACGCTCTTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAGAGCCTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACACAGAAGAGGAAGTGAACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACTCACTTCGCTTTACAGGT
CTCCATTTGTGTTAAAAAAGGTTTTAGGCCAGGCGCGGTGCTCAGCCTGTAATCCAACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTGAGAGCACGCTGGCCAACTAGGTG
AAACCTCCGTCTCTATTAAAAATACAAAAATAGCCGAGAGTGGTGGCATGCACCTGTCAATCCCA
GCTACTCGGAGGCTGAGGAGGAGAAATCGCTTGAACCGGAGGCGAGGTTGCAGTGAAGCCGA
GATCGGCCACTGCATCCAACTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAGATTTTATTAAAGATATTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NNVLALGQCVLGAFASFYWFHKKPDITFPLISAFIRTLRYHTGSLAFGALILTIVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGNFCVSAKNAFMLMRN
IVRVVVLDKVIDLLLFPGKLLVVGGVGLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNGSLDRPTYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

[illegible]

1

FIGURE 12

MGACLGACSLSCASCILCGSAPCILCSCCPASRNSTVSRLIFTFFFLGLVLSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFTAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLFILIQLVLLIDFAHSWNQRNLGKAE
ECDSRAWYAGLFFFTLLFVLLSIAAVALMFMYTEPSGCHEGKVFIISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRRMISTNTAVVWKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGGCCGGCCAGGAACCCGTTAAAGGTGCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAAACCTCCTCCAGGATGAACCCACTGCCAGAAGACATGGAGAAGC
CTCTACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCGACACAA
CTCATGGCCAGGATTGAGTCTTATGAAGGAAGGGAAAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGGTTGTACACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTTCTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAAGTGCCTTTTACTAGCAAAAG
TGATCCTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCAITCATC
CTTGCCCTGGATTGAGACGTGGTTCTTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCCCTCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCGAGGAGTGGAGTCTCCCTGTGCACAGTAAAGTTGAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCCTGCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAATAGGTGAAGAAAGTCTGTGCTGTATTCTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTAC
AGAATGGAAATTTTTTGTTCATGCTCAGATTATTTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACACTCATGCACATGIGCTCTTTGTACAGTTTTTAAAGTGTAAATAAATCTG
ACATGTCATGTGGCTAGTTTTATTTTCTGTTTTGCAATTATGTGTATGGCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATATACGTAAATGTCCACGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATGTTCTGAAATGTTCTAAATCTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTCATATCATGTTTCCTTTGCGCTCAGCCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVREFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLPPQAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTGGAACGAGTGTGCTTCGGGACCCAGGACCCCCCTGCGGCCGACCCCGCAGGAAGAGTGGG
 CGCGGCGCTGCCCCCGCGGCTCCCTGGCGCGCGCGCGCTCCCGGGACAGAAGAATGCTGCTCAG
 GGTCCCTCTGCTGCTGCGCGCTGCTCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCAT
 CGCGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCAGGGGACCCAGGTGCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAAGACGGCATCACCATGCTCGAGCG
 AGGCAGCTTTGCGGCGCTGCGGGCGCTGAGCTCTGGACCTGTACAGAACAGATGCCAGCG
 TGCCAGCGGGGTCTTCAGGCACTCGCCAACTCAGCAACTGGACCTGACGGCCAAAGGGTG
 CATGAATCACCAATGAGACCTTCGCTGGCTGCGGCGCTCGAGCGCTCTACTGGGCAAGAA
 CCGCATCGCCACATCCAGCTGGTGCCTTCGACAGCTCGACCGCTCTGGAGCTCAAGTGC
 AGGACAACGAGTGGGGCACTGCCCGCGCTGCGCTGCCCGCTGCTGCTGCTGGAACCTCAGC
 CACACAGCTCTCTGGCCCTGGAGCCGCGCTCTGCACTGCCAAGTGGAGCGCTGCGGCT
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACGCGCTTGCGCAACTCCAGACC
 TGGATGTGTCCGACACAGCTGGAGCGAGTGCCACTGTGTATCGAGGCTCCGGGGCTGAGC
 CGCTCGGCTGGCGCGCAACACCGCATTTGCCAGCTGCGGCCGAGGACCTGGCGGCTGGC
 TGCCCTGCAGGAGCTGGATGTGAGCAACTTAAGCTGCAGGCGCTGCTGGCGACCTCTCGGGC
 TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGC
 TGGTTTGGCCCTGGGTGCGGAGAGCCAGTCACTAGGCCAGCCCTGAGGAGACGCGCTGCCA
 CTTCCCGCCCAAGAAGCTGGCGGCTGCTCTGGAGCTTGAATACGCCACTTTGGCTGCCAG
 CCACCACCACACAGCCACAGTCCCAACACGAGGCCGTGGTGGGGAGGCCACAGCTTGTCT
 TCTAGCTTGGCTCTTACTGGCTTAGCCACAGCGCGCGCACTGAGGCCCCAGCCCGCGCTC
 CACTGCCCCAGCGACTGTAGGCGCTGTCCCCAGGCCACAGGACTGCCACCGTCCACCTGCCCA
 ATGGGGGACATGCCACCTGGGACACGGCACCTGGCGTCTTGTGCCGGAAGCTTCAGC
 GGCTGTACTGTGAGAGCCAGATGGGGAGGGGACAGGCCCAAGCCCTACACAGTCAAGCCGAG
 GCCACCACGGTCCCTGACCTGGGCATCAGCCGGTGAGCCCACTCCCTGCGCGTGGGGCTGC
 AGCGTACTCTCAGGGAGCTCCGTGCACTCAGGAGCTCGCTCCTACCTATCGCAACCTATCG
 GGCCCTGATAAGCGGTGGTGACCTGCGACTGCTGCTCGCTCGCTGAGTACAGGCTCACCA
 GCTGCGGCCCAAGCCACTTACTCGCTGTGTATGCTTTGGGGCCGGGCGGGTCCGGAGG
 GCGAGAGGCGCTGCGGGAGGCCATACACCCCAAGCCGCTCACTCCAACACGCCCCAGTCAAC
 CAGGCCCGCAGGGCAACTGCCGCTCTCATTTGCGGCCCGCTGGCGCGGTGCTCTCGGGCGC
 GTGGCTGCGGTGGGGGAGCTTACTGTGTGCGCGGGGCGGGCATGGCAGCAGCGGCTCAGG
 ACAAGGGCAGTGGGGCAGGGGCTGGGCCCTTGAAGTGTGAGGAGTGAAGTCTCCCTTGGAG
 CCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGCCCTGCCAGCGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTCCAGGGCTGGGCTCCAGTCAACCTCCAGCGCAAGCCCTACACTTAAGCCA
 GAGAGACAGGGCAGCTGGGGCGGGCTCTCAGCAGTGAAGTGGCAGCCCCCTCTGCTGCTGC
 ACACACAGTAAAGTTCTCAGTCCCAACTCGGGATGTGTGAGCAGAGGGCTGTGTACCAAGCT
 GGGCGCTGTTCCTCTGGACCTCGGTCTCTCATCTGTGAGATGCTGTGGCCAGCTGAGAGCC
 CTAAGTCTCCCAAGAACGATGCTTATGAGGACAGTTCGCGCTGCTCCGCAACAGTGCAGT
 CTGGGACGGCGGGGCTGCCATGTGCTGCTTAAGCATGCTGGGTCTGCTGGGCTCTCCAC
 TCCAGCGGACCTGGGGCCAGTGAAGGAGCTCCCGGAAGAGCAGAGGAGAGCGGCTAGC
 GGCTGTGTACTCTAGTCTTGGCCCCAGAGCGAAGGAGCAAGAACTGGAAGAGAGATGCT
 TTTAGACATGTTTTGCTTTTTTAAATATATATTTTATAGAGATCTTTCCATTATTTCT
 GGGAGATGTTTTTCAACTCAGAGACAGGACTTTGGTTTTTGAAGCAACAGATGATATGAA
 GGCCTTTTGAAGAAAAATAAAGATGAAGTGTGAA

FIGURE 16

MCSRVPILLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPFDTVGLYVFENGIT
MLDAGSFAGLPGQLLDLSQNIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRLRLRLY
LGKNIRIRHIQPGAFDTLDRLLEKLQDNELRALPPLRLPLRLLDLSHNSLLALEPGILDTANVE
ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPVIRGLRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPLRLLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWSPTAPATEAP
SPPSTAPPTVGVFPQDCPPSTCLNGGTCHLGRHHLACLCEGFTGLYCYSQMGQGTTPSPPT
VTPRPRLSLTLGIEPVSPSTSLRVGLQRYLQSSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRGRAMAAAAQDKGVGPGAGPLEGVRKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFFPGPLQSPHLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTATGCGG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGGCGGTGCTGCTGAGTGTGGCTCGGCGTCTCTCGG
ATGAAGAAGGCGACCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCGAGCAGAGTAGTTGCTGCTCAAAATATTTCTTGATTCAGAAGATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAAAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCAGATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGCCAGACTGTGGTGTG
CTACAACCTATGACTACAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATCGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAGAGAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAGCTGACTGAGGAAGGCTCTCCCAAGGCACAGACTGCTTTGG
CTTCTGTATGCCCTCGGACTTTGGTGTAAATCAAGTCAGGCAAAAGGCTCTTGATATATTATACAT
TTGGAGCTCTTTGGGGCAATCTAATAGCCACATGTTTGGTAAGTAGACTTTAGTGGAGGCT
AATAATATTAAACATCAGAAGAATTTGGGTTTATAGCGGCCACAACTTTTTCAGCTTTTCATGATC
CAGATTTGCTTGTATTAAAGACCAATATTCAAGTTGAATTCCTTCAAATCTTGTTAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTTGGTGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGCTGATTGCCCTTAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATGTGAGAATCATTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTAGTATTTTCA
AAATTTGTAAATGGTGCCTATAGAAAAACAACATGAATATATACAATATTTGCAACAATGC
CCTAAGATTGTTAAATTCATGGAGTTATTTGTGCAGAAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTCCCAATTTATCTGGTCATTTATTTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGTCAATTTGTACTTTTCTTTGCTAATTTGG
AAGATTAACTCATTTTAAATAAATATATGTCTAAGATTAAAAAAGAGAGAGAGAGAGAGAGAGAG
AAAAAAG

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSEVKDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSGEGESVTEDISFLESNPNKDYEEPKVKRKPALTAIEGTAHGEPCHPFFLFLLDK
EYDECTSDGREDGRLWCATTYDYKADEKWFCEEEEEAAKRRQMGEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFCALGGNLIAMVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTTCAGATTTTAAAGCCATTCTGCGATGGAATTCATGAAC TAGCAAGGAGACCATCTCTTT
 GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
 CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAGGAAAACTAAAGATTGAAGACATCACTG
 ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
 AAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
 AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
 CCCAGTGGGTGAAGAACCAGTTGGGGAGAAAGTCTCTGGGGTCTGATCAATAATGCTGCTGTT
 CCCGGCTGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
 CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCTTTGGTCAAGAAAGCTCAAGGGAGAG
 TTATTAAATGCTCCAGTGTGGAGGTGCGCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
 TATGCGATGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGACGCTCTC
 ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
 TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAACCAACAATATGGAGAAGGTTACATTGAAAAA
 AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
 CATGGACCACGCTTAACAAGTCTCTTCCCTAAGACTCATTATGCGCTGGAAAAGATGCCAAAA
 TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTATTGTTGAAACAGAAA
 GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCAACAATGTCTCTCCAGGCTATGA
 AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
 CATTTAGATCGTCTTATTGATTGCAAAAGGGAGTCCACCATCGCTGGTGGTATCCAGGGT
 CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
 GTATTTAGGCTTTGCTGCTTGGTGTGATGTAGGGAAATTGAAAGACTTGCCCATTCAAATGA
 TCTTTACCGTGGCCTGCCCATGCTTATGCTCCAGCATTACAGTAAC TTGTGAATGTTAAGT
 ATCATCTCTTATCTAAATATTAAGATAAGTCAACCCAAAAA AAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAA.

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCEIPEGLFKTNLADPVKVIKKLAIWEQLSPDIKQOYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAEANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGAGAGTACGTCGCGGTCCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACTCAACACGGACTCGGACACGGAAGSTTTCTCTTGGGGAAGTAAG
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTTCAGAAATATATTCATGCTATCAGCTTTTAGCTTTTATRAATTCTCAGGCGAAGTAATG
 AGCAAGCACTGAAGAAATATTATCAAAATGTCAAAAGAATGTGGTAGGTTGGTACAAAATCCGT
 CGTCATTGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAAGGACATTT
 TTCAAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCTTTATATAAACTCAAAAGGACTTTTTACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGCTGAACAACTGGGTTATAAACTGTATCAGGTTCTGTATGCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTGAGGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
 ACATTTTTCTTTGTGAGGCAATACGACCTTTTTTCCAAATCTGAATTTCTTCATTATGTTGTT
 ATGCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACTACAACCCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACATGACATTCCTGAAGCTAGTCCAGCTAGTACACCCAC
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGSTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
 GAGCAGCCGAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAAATATTCACGGTCTC
 CTACATTTTGAATCCTTTTAACTTACAAAGGAGATTTTTTATTGCTGATGGGTAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTGCAAGTAAGTTCAITTTGTTTTACTATGTTCAACC
 TGTTTGAGTAATACACAGATAACTCTTAGTGCAATTACTTCACAAAGTACTTTTTCAAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTCACTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATCTTTAGAATTGGAAGATGAGACCGACAGTGCGCTCACACCTGTAATCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGCTATTAATAAATAAATGGAAGCAAGATAGCCTTATTTTCAAAATATGGAAA
 GAAATTTATATGAAATTTATCTGAGTCATTAAAATCTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCAATAAATTGCAAAACATCATCT
 AAAATTTAAAAAATAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLEVKGEAKNSITDSQMDDEVVYITIDIQKYI
PCYQLFSFYNSSGEVNEQALKILSNVKKNVGVWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIIITESCSTHRLEHSLYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFFEEDGSLKEVHKINEMYASLQELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCYNHHLDVVDNLTL
MVEHTDIPASFPSTPQIIKHKALDLDLRWQFKRSRLDLDQDKRSKANTGSSNQDKASKMSSPET
DEEIERMKGFGEYSRPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGGCGGAGGGCAGAGTCAGCCGAGCGAGTCAGCCGGACGAGCGGACCGAGCGCAGGGCAGCCCAA
GCAGCGGGCAGCAACGCCCGCGGCCCCACACCCCTCTCGGCTCCCGCGGGGCTTSCACCCCTTCCTCTCTCC
GGCTCCCGCGCTCGCCGGCAGTCAGCTTGCGGGTTCCGCTGCCCGCGAAACCCCGAGGTCACAGCCCGCGGCTCT
GCTTCCTCGGCGCGCGGCCCTCCACGCCCTCTTCTCCCTGCGCCCGGCGCTGGCACGCGGGACCGTTGCTGA
CGGAGGCGCAGCTCTACTTTTCCGCCCGCGCTCTCTCGCTGCTCGCTCTTCACCAACTCCAACCTCTCTCC
TCAGCTCCACTCGCTAGTCCCGGACTCGCGCAGCCCTCGCGCGCTGCCGTAGCGCGGCTTCGCGTCCGCTCCAAA
GGTGGGAGCGCTCCGCCCGCGCGCCAGCAATGGCACGGTTCCGCTTGC CGCGCTTCTCTGCACTTGGCATGCTC
AGCGCGCGGCTGCTGCTGCGCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGGAGCTTTACGTGTCCAAAGGCTTC
AACAGAACGATGCCGCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGCCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGATGACGCTGCAAGTGAAGATGATTTCAAAGTGTGTGAGCAACAGTGCATCATTTG
CAAGCTGTCTTCTTCCACTTACAGAGAGTTTGATGAATCTTCAAAGACTACTTTGAAATGCGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGCCAATTTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTCTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTCTGGGCTCGCTCTCGAGCGGATG
TTCCGCTGTGTGAATCCCATGACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAG
CCCTTCGGAGATGCTCCCTGCAAAATGAAGCTCCAGGTTACTCGTCTTTGTAGCAGCCGCTACTTTGCTCTAAGG
TTAGCGGTTCCGGGAGATGCTGCTGAGCAAGGCTCCCGTGAACCCACAGCCAGTGTACCAATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCGCGGGCTCTCGTACTGTGAAGCCATGTTACAACTACTGCTCAAAATCATGAGAGG
TGTTTGGCAACCAAGGGATCTCGATTTGAATGGAACAATTCATAGATGCTATGCTGATGTTGGCAGAGGCTA
GAGGCTCTTTCAACATGAATCGGTATGGATCCCATGGATGTGAAGATTTCTGATGCTATATGAACATGACGAT
AATAGTGTCAAGTGTCTCAGAGGTTTTCAGGAGTGTGAGCCGCCAAGCCCTCCAGCTGGACGAATTTCTGCT
TCCATCTCTGAAGTGCTTCACTGCTCGCTTCAGACCACATCAOCCCGAGGAAAGCCCAACACAGACAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAATGAAACAGGCCAAGAAATTTCTGCTCCCTCCCGAGCAGC
GTTTGCAACGATGAGAGGATGGCTGAGGAAACGGCAATGAGGATGACTGTTGGAATGGAAAGCCAAAGCAGGATC
CTGTTTGCAATGACAGGAAATGGATTAGCCAACAGGGCAACACCCAGAGGCTCAGGTTGACACCAAGCAACAGC
ATACTGATCCTGCTCAAATCATGGCTCTTCGAGTGATGACCAAGATGAAGATGCATCAATGGGAGCAGCTG
GACTCTTTGATATCAGTGATGAAAGTAGTGAGAGGAAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACATGCACTGACGATTTTATATGTGGCCACTGGTTTGAAGTGTGCTGCTTTGTTCTCATTCAGTTTGGG
AGGAAAGGAGCTGTGCATTTGAGTTGCTGCTCCCCAAACATGTAAACCTGGCTACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTTGTGATTTATCACTATTAITTTGTTGATGTTTCTCATTTGTTGCGGTT
TTTTTTTCCAACTGTGATCTCGCTTTGTTCTTACAGCAACAGGCTCCCTTTCTGGCAGCTAACATGACGTATT
TCTGAATATTAATAGCTGTACAGAGCAGGTTTATTTATCATGTATCTATTAAAGAAAGGCCAAAGAGC

FIGURE 24

MARFGLFALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDPKSVVSEQC�HLQAVFASRYKKFDEFFKELLENAEKSINDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPPGDVFRKLQVTRAFVAARTFAQGLAVAGDVVSKSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCINMRGCLANQGDLD FEWNNFIDAMLVVAERLEGPPNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFGCGGPPKPLPAGRISRISISESAFSARFRPHHPPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSSLPSNVNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDTIS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGGCEYQQCFSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTGCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCAACAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAGATTGGTTCTTGAG
AGCCCCGAGAAGAAATTCATGACAGTGTCTGGGCTGCCAAGAAGCAGTGCCCCGTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGCAGTGAGCACACCTACCAGACACTC
TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCAGTGTCTCAGAAAGCA
TGTTTTTCAAGATCATTTTGTGTTGCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAATCTGAGCTTCCT
AGCTAGTGTCAATTAACCTTAATGCAATCAGGAAGTAGCAACAGAAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAAA

FIGURE 27

GGAGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTSCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCGATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTCTGCTTTTACCTGCG
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCTGTGCAAAAGCTACTTCCCTTACCTGA
TGCCCGTCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGAACCCAAATCCCACCTTTGAGA
AGTTCTTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGGTGGTCTGCAGCTCTGGTGGTGTG
CTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGTCCGGAGAGTACTGAGACCGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCTTTCATGTGGCAGCAAGTTTTC
GAGCCACCTGGAACACATTGGGGATGGCTGCTGCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTG
GGCCCCACATCATGGGAAGGCTGTCAAACAATCTTCCCAAGCTCCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCACCGACCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAAGACATTCAATGATACCACTACTAGTCCCTCTCCTCCCAACCTCTGCCAGGGC
AATCTCTAATCTCAATCCCGCCTTCGACAGTGAAAAGCTCTACTTCTACGTGACCCAGGGAGG
AAACACTAGGACCTGTGTATCTCTCAACTGCAAGTTTCTGGACTAGTCTCCCAAGCTTTGCTC
CCAATGTTGCCCTTTCCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCACTACCACCTTCTTCTGAGCTGGGGCACCAGGGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTTTAATAAATAGACGAA
ACCCAG

FIGURE 28

MDILVPLIQLLVLLLTLP LHMALGCGWQPLCKSYFFYLMAVLTPKSNRRMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDEVVCTLVLCVQSPRKVLQEVRRVLRPGGVLPFWEHVAEYPYGSWAFMWQQVFETW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLVPVPHMGKAVKQSFSSKALICSFPSL
QLEQATHQFIYLP LRG T

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGGCTTCTTTCCCCCTTCTTTCTGTGTCTCCTGCCTCAT
CGGCCCTGCCATGACCTGCAGCCAAGCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGATGGCTA
AGAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTTGTAAGATTGTCTGTGTAAATATGCTTTATAATAAAC
AGTTAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPTRDWAFPPPSFLCLLPHPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

10063713

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCAGTGTGCCAGGCTACAGTT
 CCTCCAAGCAAGTCATTCCCTTATTTAACCGATGTGCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCTCCACCGAATTCTAAGTGAATCAGTCCG
 GAAGAGATACAATCCTTGGCTGTGTATCCTCGCATAGCCTTGTCTTTGGCCATGATGTTTACC
 TTCAGATTTCATCACCACCTTCTGGTTTACATTTTCATTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGAAAAATATGAAGTGGCTGTGGGTTTGCTATCGTATCCACAGGCATCAGGCGATG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTGTAGCTTTTCCAAATCAC
 AAATAAGCCATCAGCAGTGTCTCCCTTCTGTCTGTTCCAGCCACTGTGGACATTGCCATCTCTCA
 TTTTCTCTGGGTCTCTGGTGGCTGTGCTGCTGAGCCTGGGAACCTGCAGGAGCTGCCAGGTT
 ATGGAAGGGGCCAAGTGAATATAAGCCCTTTGGGCAATCGGTACATGTGGTCTGTAACATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGGCTGCCAGCAATGACTATAGCTGGGCGAG
 TGGTTACTTGTATTTCACAGAAGTAAAAATGATCCTCTGATCATCCATCCTTTTCTCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTGTGAAGGGTCACTTTTAACTCTGTGGTGAG
 GATTCCGAGAACTATTGTCTGTACATGCAAAACGCACTGAAAGACAGCAGCATGGTGCATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACAGAATGCATATACTACAACCTGCTATTAAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCCAAGAACTCAAGTCACCTTACATCTATTAACCTGCTTTGGAGACTTCATAA
 TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGAGGACTCATGGCTTTTAACATC
 AATCGGGCATTCCAGGTGTGGGCACTCCCTCTGTATTGGTAGCTTTTTTGGCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAACCTGTGCTGGATGCACCTTTTCTGTGTTTTGCTGTGATC
 TGGAAACAAATGATGGATCGTCAGAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGAGCAACAAATTAACAATGCAAGGCGACAGCAGGACAAGCACTCATTAAGGAATGAGGA
 GAGAACAGAACCTCAGGCCATTGTGAGTAGATACCCATTTAGGTATCTGTACCTGAAACCAATT
 TCCTTCTAAGAGCCATTTACAGAATAGNAGATGAGACCCTAGAGAAAAGTTAGTGAATTTTTT
 TTAAGAGCCTAATAAACCTATTCTTCTCAAAA

FIGURE 32

MSGRTILGLCILALALS LAMMTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTDLSIE
LDTARENMKCVLGFIVSTGITAVLLVLI FVLRKRIKLTVELFOITNKAISSAFFLLFQPLWTF
ILIFFVVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQSTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FII FLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

1. $\frac{1}{2}$ 2. $\frac{1}{3}$ 3. $\frac{1}{4}$ 4. $\frac{1}{5}$ 5. $\frac{1}{6}$ 6. $\frac{1}{7}$ 7. $\frac{1}{8}$ 8. $\frac{1}{9}$ 9. $\frac{1}{10}$ 10. $\frac{1}{11}$ 11. $\frac{1}{12}$ 12. $\frac{1}{13}$ 13. $\frac{1}{14}$ 14. $\frac{1}{15}$ 15. $\frac{1}{16}$ 16. $\frac{1}{17}$ 17. $\frac{1}{18}$ 18. $\frac{1}{19}$ 19. $\frac{1}{20}$ 20. $\frac{1}{21}$ 21. $\frac{1}{22}$ 22. $\frac{1}{23}$ 23. $\frac{1}{24}$ 24. $\frac{1}{25}$ 25. $\frac{1}{26}$ 26. $\frac{1}{27}$ 27. $\frac{1}{28}$ 28. $\frac{1}{29}$ 29. $\frac{1}{30}$ 30. $\frac{1}{31}$ 31. $\frac{1}{32}$ 32. $\frac{1}{33}$ 33. $\frac{1}{34}$ 34. $\frac{1}{35}$ 35. $\frac{1}{36}$ 36. $\frac{1}{37}$ 37. $\frac{1}{38}$ 38. $\frac{1}{39}$ 39. $\frac{1}{40}$ 40. $\frac{1}{41}$ 41. $\frac{1}{42}$ 42. $\frac{1}{43}$ 43. $\frac{1}{44}$ 44. $\frac{1}{45}$ 45. $\frac{1}{46}$ 46. $\frac{1}{47}$ 47. $\frac{1}{48}$ 48. $\frac{1}{49}$ 49. $\frac{1}{50}$ 50. $\frac{1}{51}$ 51. $\frac{1}{52}$ 52. $\frac{1}{53}$ 53. $\frac{1}{54}$ 54. $\frac{1}{55}$ 55. $\frac{1}{56}$ 56. $\frac{1}{57}$ 57. $\frac{1}{58}$ 58. $\frac{1}{59}$ 59. $\frac{1}{60}$ 60. $\frac{1}{61}$ 61. $\frac{1}{62}$ 62. $\frac{1}{63}$ 63. $\frac{1}{64}$ 64. $\frac{1}{65}$ 65. $\frac{1}{66}$ 66. $\frac{1}{67}$ 67. $\frac{1}{68}$ 68. $\frac{1}{69}$ 69. $\frac{1}{70}$ 70. $\frac{1}{71}$ 71. $\frac{1}{72}$ 72. $\frac{1}{73}$ 73. $\frac{1}{74}$ 74. $\frac{1}{75}$ 75. $\frac{1}{76}$ 76. $\frac{1}{77}$ 77. $\frac{1}{78}$ 78. $\frac{1}{79}$ 79. $\frac{1}{80}$ 80. $\frac{1}{81}$ 81. $\frac{1}{82}$ 82. $\frac{1}{83}$ 83. $\frac{1}{84}$ 84. $\frac{1}{85}$ 85. $\frac{1}{86}$ 86. $\frac{1}{87}$ 87. $\frac{1}{88}$ 88. $\frac{1}{89}$ 89. $\frac{1}{90}$ 90. $\frac{1}{91}$ 91. $\frac{1}{92}$ 92. $\frac{1}{93}$ 93. $\frac{1}{94}$ 94. $\frac{1}{95}$ 95. $\frac{1}{96}$ 96. $\frac{1}{97}$ 97. $\frac{1}{98}$ 98. $\frac{1}{99}$ 99. $\frac{1}{100}$ 100. $\frac{1}{101}$ 101. $\frac{1}{102}$ 102. $\frac{1}{103}$ 103. $\frac{1}{104}$ 104. $\frac{1}{105}$ 105. $\frac{1}{106}$ 106. $\frac{1}{107}$ 107. $\frac{1}{108}$ 108. $\frac{1}{109}$ 109. $\frac{1}{110}$ 110. $\frac{1}{111}$ 111. $\frac{1}{112}$ 112. $\frac{1}{113}$ 113. $\frac{1}{114}$ 114. $\frac{1}{115}$ 115. $\frac{1}{116}$ 116. $\frac{1}{117}$ 117. $\frac{1}{118}$ 118. $\frac{1}{119}$ 119. $\frac{1}{120}$ 120. $\frac{1}{121}$ 121. $\frac{1}{122}$ 122. $\frac{1}{123}$ 123. $\frac{1}{124}$ 124. $\frac{1}{125}$ 125. $\frac{1}{126}$ 126. $\frac{1}{127}$ 127. $\frac{1}{128}$ 128. $\frac{1}{129}$ 129. $\frac{1}{130}$ 130. $\frac{1}{131}$ 131. $\frac{1}{132}$ 132. $\frac{1}{133}$ 133. $\frac{1}{134}$ 134. $\frac{1}{135}$ 135. $\frac{1}{136}$ 136. $\frac{1}{137}$ 137. $\frac{1}{138}$ 138. $\frac{1}{139}$ 139. $\frac{1}{140}$ 140. $\frac{1}{141}$ 141. $\frac{1}{142}$ 142. $\frac{1}{143}$ 143. $\frac{1}{144}$ 144. $\frac{1}{145}$ 145. $\frac{1}{146}$ 146. $\frac{1}{147}$ 147. $\frac{1}{148}$ 148. $\frac{1}{149}$ 149. $\frac{1}{150}$ 150. $\frac{1}{151}$ 151. $\frac{1}{152}$ 152. $\frac{1}{153}$ 153. $\frac{1}{154}$ 154. $\frac{1}{155}$ 155. $\frac{1}{156}$ 156. $\frac{1}{157}$ 157. $\frac{1}{158}$ 158. $\frac{1}{159}$ 159. $\frac{1}{160}$ 160. $\frac{1}{161}$ 161. $\frac{1}{162}$ 162. $\frac{1}{163}$ 163. $\frac{1}{164}$ 164. $\frac{1}{165}$ 165. $\frac{1}{166}$ 166. $\frac{1}{167}$ 167. $\frac{1}{168}$ 168. $\frac{1}{169}$ 169. $\frac{1}{170}$ 170. $\frac{1}{171}$ 171. $\frac{1}{172}$ 172. $\frac{1}{173}$ 173. $\frac{1}{174}$ 174. $\frac{1}{175}$ 175. $\frac{1}{176}$ 176. $\frac{1}{177}$ 177. $\frac{1}{178}$ 178. $\frac{1}{179}$ 179. $\frac{1}{180}$ 180. $\frac{1}{181}$ 181. $\frac{1}{182}$ 182. $\frac{1}{183}$ 183. $\frac{1}{184}$ 184. $\frac{1}{185}$ 185. $\frac{1}{186}$ 186. $\frac{1}{187}$ 187. $\frac{1}{188}$ 188. $\frac{1}{189}$ 189. $\frac{1}{190}$ 190. $\frac{1}{191}$ 191. $\frac{1}{192}$ 192. $\frac{1}{193}$ 193. $\frac{1}{194}$ 194. $\frac{1}{195}$ 195. $\frac{1}{196}$ 196. $\frac{1}{197}$ 197. $\frac{1}{198}$ 198. $\frac{1}{199}$ 199. $\frac{1}{200}$ 200. $\frac{1}{201}$ 201. $\frac{1}{202}$ 202. $\frac{1}{203}$ 203. $\frac{1}{204}$ 204. $\frac{1}{205}$ 205. $\frac{1}{206}$ 206. $\frac{1}{207}$ 207. $\frac{1}{208}$ 208. $\frac{1}{209}$ 209. $\frac{1}{210}$ 210. $\frac{1}{211}$ 211. $\frac{1}{212}$ 212. $\frac{1}{213}$ 213. $\frac{1}{214}$ 214. $\frac{1}{215}$ 215. $\frac{1}{216}$ 216. $\frac{1}{217}$ 217. $\frac{1}{218}$ 218. $\frac{1}{219}$ 219. $\frac{1}{220}$ 220. $\frac{1}{221}$ 221. $\frac{1}{222}$ 222. $\frac{1}{223}$ 223. $\frac{1}{224}$ 224. $\frac{1}{225}$ 225. $\frac{1}{226}$ 226. $\frac{1}{227}$ 227. $\frac{1}{228}$ 228. $\frac{1}{229}$ 229. $\frac{1}{230}$ 230. $\frac{1}{231}$ 231. $\frac{1}{232}$ 232. $\frac{1}{233}$ 233. $\frac{1}{234}$ 234. $\frac{1}{235}$ 235. $\frac{1}{236}$ 236. $\frac{1}{237}$ 237. $\frac{1}{238}$ 238. $\frac{1}{239}$ 239. $\frac{1}{240}$ 240

Page 187 of 322

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCVDKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSSSKSPAQAQGETTKAYQRPPIPGTTAQPVTLMLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLFVSLGDPNCKIDLFLIDGSTSIGKRRFRIQRQLLADVQAALDI
GPAGPLMGVVQYGDNPATHFNLKTHNSRDLKTAIEKITQRGGLSNVGRAISFVTNFFSKANGN
RSGAPNVVVMDVGWPTDKVEEASRLARESGINIFFITIEGAANEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLFEGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAMAAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

1

FIGURE 36

MASALMTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNMWYFYEYEPYRQDFHF
TLREHSNCSHQNFLLVILVTSHPSDVKARQAIRVTWGEKKSWSWGYEVLTFLLGQEA EKDKMLA
LSLEDEHLLYGDIIROQDFLDYNNLT LKTIMAFRWWTEFCFNKYVMKTD TDVFINTGNLVKYLL
NLNHSEKFFTGYPIDNYSYRGFYQKTHISYQYEPKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KP IKFEDVYVGICLNLKVNIIHPEDTNLFFLYRIHL DVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Lx-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCAGCAGCCGCGGCAAGGATGGAGCTGGGTGTGGAGCGAGTTGGGGCTCACTTTCTCAGCTCCTTCATC
TCGTCTTCCGAAGAGATACACAGTCATTAATGAAGGCTGCGCTGGAGCAGAGTGGAATATCATGTGTGCGGAGTGTCTGTG
AATATGATCAGATTGAGTGCCTTCCGCCGGAAGAGGGAAGTGTGGGTATACCATCCCTTGTGCGAGGAATGAGGAGAA
TGA GTGTGATCTCTGCTGATCCACCGAGTTGTACCATCTTTGAAAAC TCGAAGAGCTGCGGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAGGGGTTTCTACTGTGCGAGGTGCCGAGCAGGCTGTACGAGGAGAGACTGCATGCGATGTG
GCCAGTGTGCGAGCCCCAAGGGGTGAGATTGTGTGGAAAGCTATCCCTAAATGCTCAGTGTGAATGGAGCAATCATGC
TAAAGCTGGGTTTGTATCCAACTAAGATTGTGTCAATGTGAGCTGGAGTTTGAATACATGTGCCAGTATGACTATGTGAG
GTCTGTGATGGAGACAACCGGATGGCCAGATCATCAGCGGTCTGTCTGGCAACGAGCGCCAGCTCCTATCCAGAGCATAG
GATCTCTCACTCACGTCTCTTCCACTCGGATGGCTCCAGAAATTTTGAAGGTTTCCATGCAATTTATGAGGAGATACAGC
ATGCTCTCATCCGCTTGTTTTCATGACGGCACGTGCTCCTTGAAGAGGCTGGATTTACAGTGTGTGCTGTGTGCAAGC
TATATCTGGCAGCGCTGTGAAATCTCTTTGAAGAAAGAACTGCTGAGACCTGTGGGGCCAGTCAATGGGTACAGAA
TAAAGGGGGCCCTGGGCTTATCAACGAGCGCCATGCTAAATTTGGCACCGTGGTGTCTTTCTTTGTATCAACTCTTATGT
TCTTATGGCAATGAGAAAGAAC TTGCCAGCAGAAATGGAGAGTGGTCAAGGAAACAGCCCATCTGCATAAAGGCTTGCCGA
GAACCAAGATTTCAGACCTGTGAGAAAGAGATTCTCCGATGAGGTTCACTCAAGGAGAGACCAATTACACAGCTAT
ACTCAGCGGCTTTCAGCAGCAGAACTGCGAGAGTGGCCCTTACCAAGAAAGCCAGCCCTTCCCTTTGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCCCTTCTACGCGCCCTGGGACAGCAGGAGGACATGT
CTGAGAGCTGGGAAGTGGAGTGGGGGGGACCATCTCTGCATCCCTATCTGCGGGAAATTTGAGAACATCACTGCTGTCAAAAG
CCCAAGGTTGCTGCTGCGCTGGCAGGACCATCTACAGAGGACCGCGGGGTGATGACGAGCAGCTACACAGAGGAGAG
GTGCTCTCTAGTGTGACGCGGTGCGCTGTGGAATGAGGCACTGTGGTGTGGTGTCCCATCTGTGTACTGTGGGAGAG
GTACCATGATCAAGACAGCAGACTGAAAGTTGTTTGGGGAATTCATACCGGATGATGACGGGATGAGAAAGACCTCC
AGAGCTTACAGATTTCGTATCAATTCATGATCCCACTATGACCCCATCTGCTTGTGATGTGACATGCCATCTCGAAGCT
CCTAGACAAAGGCCGTATCAGACCCGAGTCCAGCCCATCTGCTGCTGCGCACTCGGATCTCAGCACTTCTCTCAGAGAG
TCCACATCACTGTGCTGTGCTGGAATGCTGCGAGAGCTGAGGAGGCTTGGCTTCAAGAACGACACTGCGCTCTGGGG
TGTGCAGTGTGGTGAATCTGCTGCTGTGTGAGGAGCAGCATGAGGACATGGTCACTCCAGTGAATGTCACTGATACATGTT
CTGTGCGAGCTGGGAACCACTGCGCCCTTCTGATATCTGCACTGCAAGAGCAGGAGGACATGCGGCTGTGTCTTCCGGGA
CGAGCATCTCTGAGCCAGCTGGCATCTGATGGAGTGTGAGCTGAGCTATGATAAAACATGCGAGCCAGAGCTCTCCA
CTGCTTCTACCAAGGTGCTGCC TTTTAAGACTGGATTGMAAGAAATGHAAGTGAACATCATGCTCATGCACTCTCTGAGAG
TGTTCCTGTATACGCTGCTGTACGTGTGTCACTGCTGTGAAGCACTGTGGGCTGAAATGTGATTGGGCTGTGAACCTGAGT
GTGCGAGGGCTCTGACTCAGGGAJAAACTCAGTGAAGGTTGATAGACCTCATTTGCTGGTATGGCTGATGCGCGCTCCA
CTCATGAGACAGCAATGGAGATGCCAGGCTTGCAGAGATAGTTTCTTCAAGAGAGACATATACAAAGCTCTCCA
CTCCATGACTGTGTGCTTTCCCAATTTTCACTATATACAAATGCCATCAGCTTGACAGGAGAGATCTGGGCTCATGAG
GCCCTTTTGGGCTCAAGTCTTAGAGAGTGCCTGTGGGACAGCCGAGGACAGAGCTGGGATGTGGTGTGCTATGCTCT
TGTGTATGAGCCAGACTACAGTCTGTGCTCTTCTCTCCCATCTCTGTATACATTTTAAATAAATAGGGTGTGCTCT
GACTCTAAAAAAGAAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENWIMCRECEYDQIECVCPGKREVVGYT
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLTDDFYVKGIFYCAECRAGWYGGDCMRCGQ
 VLRAPKGQILLESYPLNAHCEWTHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGNRDGQII
 KRVCGERPAPIQSIGSSLHVLFHSDDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
 ACLAGYTGQRCEENLEERNCSDPGGPVNGYQKITGGPGLINGRAHAKIGTVVSFFCENNSYVLSGNE
 KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRRETPHLQLYSAAFSKQKLQSAPTK
 KPALPFGDLPNGYQHLLHTQLQYECISPFYRRRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP
 KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL
 KVVLGKFYRDDDRDEKTIQSLQISAILHFNYPDILLDADIAILKLDKARISTRVQPICLAASR
 DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHIMGLVSWSYDKTCSHRLSTAFKVLFPFKDWI
 ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCTCATCTCCTCATCTGAGAATCAGAGACATTAATCTTCTTACGGGCCCTGATTATTAACTGGCTTAATC
TGAAGGTTCTCAGTCAAAATCTTTTGATCTACTGATTGTGGGGCATGGCAAGTTTCTTAAAGAGCTTGGCTGG
TTTGGGCCCTTGAGCTGACAGAAGGTGGCCAGGAGAATGCAGCACACTGCTCGGAGATGAAGGCGCTCTGTGC
TGGTCTTGCTTGGCTCAGTCCCTGCTAACTACATTGACAATGTGGGCAACTGCACCTTCTGTATTGAGACTCTGTA
AAGTGGCTCCCACTACGGCCTGACCAAAGATAGGAAGAGCGCTCAAGAATGGCTGCCAGCGCTGTGCGAGCC
TCACAGCCAGGCTCCCTCCAGAGGTTTCTGACAGTGCACCATCTCTTAATGACAGACGAGCTGGCTAGACA
ACCTTGCTACGTGCTCTCGGACAGGAGCGGCGACGACATCAGCCAGTGGACTCTGGCCGAGCAACCGAATA
GGCACGGCCCTTTGAGAGATCACTATTAGGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGATGTCTTCGAA
GGACAAGAGCGGGAGTGCAGTTGCCAACATGCCGACAGGGCAGGGAATAATCTGAAAACACCACTGCCCTGAAG
TCTTTCCAAAGTTGTACCACTGATTCCAGATGGTGAAATTACGAGCATCAAGATCAATCAGTAGATCCCAAGTGA
GCCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCACTGGTCCATATCATTAACAACACATTTATCGTATGGG
TGATGGCCAGAGACGGCCGCTACTGCCAGGAGCATCATTTAAAGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCCTGGCGCAGCCCTGCCAGTGGCTGGCTGACTGTGATGCGTGAACAGAATGCCCA
CGAGGAACAATGGACAGGCCCGGATGCTACAGACCCGAGATGACAGCTTCAATGTGATTCTCAACAAGATAGCG
CCGAGGAGCAGCTTGGATAAACTGGTGGCGAAGTGGATGAGCCTGGGGTTTCACTTCAATGTGCGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCCTGTGTAGCCATCAATGGACATGATCTTGATATGGCA
GCCCAAGAAGTGGGGCTCATCTGATTGAGGCGAGTGAAGACGTGTCACTCGTCTGTGCCCGCAGGTCGGCAGC
GGAGCCCTGACATCTTTCAGGAAGCGCGCTGGAACAGCAATGGCAGCTGTGCCAGGGCCAGGGGAGGAGCAACA
CTCCCAAGCCCTCCATCTTACAAATTAATTTGTCTAGAGAGTGGTAAATTCACAAAAGACCCCGGTGAATCTCTCG
GCATGACCCTCGAGGGGGAGCATCACATAGAGATGGGATTTGCCTATCTATGTATCAAGTGTGGCCCGGAGGAG
TCATAAGCAGAGATGGAGAATAAAAAACAGGTGACATTTTGTGAATGTGGATGGGTGGAAGTGCAGAGGTGAGCC
GGAGTGAGGCAATGATTAATGAAGAAGCAATCATCTCGATAGTACTCAAGCTTTGGAAGTCAAGAGATGAGGC
CCCGAAGAGCTGCAGCAGCCGAGCAGCCCTGGACTCCAAACCAACATGGCCCAACCCAGTGACTGGTCCCATCTCT
GGGTGATGTGGCTGGAATTAACACGCTGCTGTATAACTGTAAAGTATTGATTACGAAGAACAACAGCTGGAGTCT
TGGGCTCTGCAATGAGGAGTTATGAAGAATACAATGGAAACAAACCTTTTTCATCAATCCATGTTGTGAAGGAA
CACCAGCATACAATGATGGAGAATTAGATGTGGTGAATTTCTCTTCTGCTCAATGGTAGAGTACATCAGGAATGA
TACATGCTTCTGCTGGCAAGACTGCTGAAGAATTAAGGAAGAATTAATCTTAATCTTCTTCTGGCTGGCAGCTT
TTTTTGAATCAATGATGGGTGAGGAGAAACAGAAATACAAATAGGCTAAGAGTTGAAACATATATTTATCT
TTGTGAGTTTTATATTTAAAGAAAGATACATGTAATAATGTGAGGAAGTATGATCTTAATGAAGGCAAGTT
ACACCTCAGAAAAATGATTCCAAAAAATTAAGTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTTACTTAC
AATCATTTTATATTTTCTTATCTCAATAAAAGCCCTAAACCAACTAAATGATTGATTGTATACCCCACTGAAT
CAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGATCATTAATGGCCATTTTAATTTACAGCT
AAATATTTTTTAAATGCATGCTGAGAAGCTTGCTTTTATCAACGAAGATAAATTTTTTTCAGAGTTAAA

FIGURE 40

MKALLLLVLPWLSPPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRSSQDGCPCASLTATAPS
PEVSAATISIMTDEPLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSSESLSIRLV
GGSETPLVHII IQHIYRDGVIARDGRLLPGDIILKVNMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDSSFHVLNKSPEEQLGIKLVKVDPEPGVFI FNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSRRQVRQSPDIFQEAGWNSNGSWSPG
PGEPSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPYIVISVEPGGVISRDGR
IKTDILLNVLDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTVSNPQTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCGAGCATTGTATCTTCAGTTGTCATCAAGTTGCAATCAGATTGGAAGCTCAACTTTGAAGCTTT
 CTTGCCCTGCAGTGAAGCAGAGATAGATATTATTACAGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCTACAAATTCGATTACTGTTGCTGTTGACTTTTGCCCTGACAGTGGTGGGTGGGC
 CACCAGTAACACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
 ATAAGACCTCATTTTGGGGAAGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGCTT
 GACAACCTGCTCTTCTGTCTCTCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAAATCCCAAGTGTCCAGAGGCCGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGTCGCCATCTCTGTTCCCCACCGAAGCAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACAGGCTGAAGG
 TAAAAAGTTAATCGAGCCAACTCTTGAATGTGGCTATCTAGAAGCCCTCAAGGAAGAAATTTGGG
 ACTGCTTTATATTCCAGGATGTGGACCTGGTACCCAGAGATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTAGGTTACAGTGGATATTTTGG
 GGGTGTACTGOCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGSAITTCCTAACCACTACTGGGAT
 GGGGAGSCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAAATTTCCCGGCCCTG
 CCTGAAGTGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGCAATGAGGTGAACSCAGAAG
 GATGAAGCTCTTACACCAAGTGTACAGAGTCTGGGAACAGATGGGTTGAGTAGTTGTCTTTATAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCGGTTTGGTGAATGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTATTCTTTGTTTGAATAATTTTGGCCCTAGAGACTTCAA
 ATAGTAGCACACATTAAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTCTTTTGTATTTCT
 TAGCAGAGCTCTCGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCAITTTGAT
 CATGAGGTTAAATATTGTAATATGGACTTGAAGGACTTTATATAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTGAGGACTCTGGTTGAAGGAGATTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAATAAGACTGCTGAATGCTTGAGAGAACAGAGTGTCTCTGCTCAAGGTAGAA
 AGGTACGAAGATACAACTGTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGT
 GAGAAGGCTCCACAAAAGAGGGGAGAAAAGGCAGCAATCAGGACACAGTGAACCTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAAGGCAGCAGTAGCTGAGCTGTTGAGGTTGCTGATAGC
 CTTCCAGGGGAGGACCTGCCACAGTATGCTTCCAGTGATGCCCAACAGAGAATACATTTCTCTATTAGT
 TTTTAAAGAGTTTGTGTAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAGTTTACAT
 ATTAACATAATAAATATGTCTATCAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLT FHL SYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGTKLTIN
EASTKKVELDNCPSVSFYLRGQSKLIFKPDLTLEEVOAENPKVSRGRYRPQECKALQRVAILVPH
RNRKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLAYSGYFGVTALSR EQFFKVN GFSNNYWG WGGED
DDLALRVELQRMKISRPLEFVGK YTMVFHTRDRGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCAGATCTGGG
CCGCTTGCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACACAG
ACGGGACAACTTGCAAGCTGCAACCCAGGACAGAGCTGGAGCCAGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACCCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCTCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAAATAAAATGGCTGGTCTTTTGTIT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFP
CIFCGGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

TGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
 TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
 TCCGTGTTGGGGCCGTGACTTTCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
 GACCTTCAACACACACCCCTCTTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGAACCCAAA
 ATCGTAATAGGGAGAGAGTAGACTTCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
 AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
 GGAGTACGTGCTGATGCTTACGAGCACCTGTCAAAGCCTAAAGTCAACATGGGTCTGCAGAGCA
 ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
 TATACCTGGAAGCCCTGGGGCAAGCAGCCAAATGAGTCCCATAAATGGGTCCATCCTCCCATCTC
 CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCCTTGGCAGGAACCCCTGTCAGCAGAACT
 TCTCAAGCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCTCATG
 GTCTCTCTGTGCTCTCTGTTGGTGCCCTCTCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
 GTTCTGTAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTGTGCGG
 ARACTCTTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
 AGAACAACTCTAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAGAT
 GGAAAATCCCCACTCACTGCTCAGCATGCCAGACACCAAGGCTATTTGCCTATGAGAATGTTA
 TCTAGACAGCAGTGCACCTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPCTCLTIYILWQLTGSAAAGFVKELVGSVGGAVTFPLSKVKQVDSIVMTFNTTPLVTIQP
EGGTIIIVTQNNRNRVDFPDGGYSIKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLWFLKREERQEEYIE
EKKRVDICRETPNICPHSGENTYDITPHNTRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGTGACCAATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTGAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAAGTGGAAATGTTTCTTCATCATTTCCTAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCAATGCTGATATCCATCCAGGCTCTCTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAACATCAGTGACATTTCATCCAGAATCCTTCAA
CTTGAGTGGTTTTCAATGACTCTTGTCACCTCCTACTGGTTTCAATAAACCCACAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAAACAACATAGG
CTTATCCACTTCTCAGTATTTTASGTCATTTGCTTGTGGAATCTGGAGTCTCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTCTTGCTGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTS~~C~~NGFSLLVLLLLGVVINAIP~~L~~IVSLVEEDQFSQNPISC~~F~~EW~~W~~FPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALY~~C~~MLISIQALLKGPLMCN~~S~~PSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGENKPTSNDT~~M~~ASGWRASS~~F~~HFDSEENK~~H~~R~~L~~I~~H~~FSVFLGL
LLVGILEVL~~F~~GLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

Category	Item	Score	Weight	Total
Mathematics	Mathematics	85	40%	34
	Science	78	30%	23.4
	History	92	20%	18.4
	Physical Education	88	10%	8.8
Language Arts	Reading	75	35%	26.25
	Writing	82	25%	20.5
	Speaking	88	20%	17.6
	Listening	90	20%	18.0
Social Studies	Geography	80	30%	24
	Government	75	20%	15
	History	85	25%	21.25
	Art	90	25%	22.5
Physical Education	Physical Education	88	40%	35.2
	Health	78	30%	23.4
	Music	85	20%	17
	Drama	90	10%	9
Other	Art	85	30%	25.5
	Music	78	20%	15.6
	Drama	88	25%	22
	Other	90	25%	22.5
Total				300

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSLICGGLLAIGIAAVLSGKCKYKS
SQKQHSFPVEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 51

GTGGACTCTGAGAAGCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGCAGAG
 AGGAGACCCCTGGTGGAGGAGAGACCTCTGGAGAGAGAGGGGCTGGGCAGAGATGAAGTTCAG
 GGGCCCCCTGGCCTGCCTCCTGCTGGCCCCCTGCCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATAATTGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
 GCGAAGGGGTGGGAAGGCCATTGGCAAAAGGCCGAGGGGAGCTGGCTCTAAAGTCAGTGAG
 GCCCTTGGCCAAGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
 TTGGCAGACAGGCAGAAAGATGTCATTGACACAGGAGCAGATGCTGTCCGGGCTCTTGGCAGGGG
 GTGCCTGGCCACAGTGGTGTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
 GAAACTCAGCAGGCAGCTTTTGAATGAATCCTCAGGGAGCTCCTTGGGCTCAAGGAGGCAATGGA
 GGGCCACCAAACTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG
 AGCCAGCAACCAGAAATGAAGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
 ACTCTGGGGAGGCAGCGGCTCAGAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAAACAAC
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGTGGACAGCGGCAGTGAGT
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCGGGCAACACGGTGGGAGCGGCGGAGGAATGGA
 CATAAACCCGGGTGTGAAAGCCAGGGAATGAAGCCCGGGGAGCGGGGAATCTGGGATTCAGGG
 CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAGAGGGCAATGCCTCCTTG
 GAGGCTCTGGAGCAATTAACGGGGCAAGGGTCGAGCTGGGGCAGTGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACITGTAACCTCTGAGACGCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAACAGGACCCAGAGAACTCTC
 GCATCCCGTGAACCTCAGACAGGAGGCCACGATTGGATGGGAGCCGCCACACTCCCTCCTTAA
 AACACCACTCTCTCATCACTAATCTCAGCCCTTGCCCTTGAATAAACCCTAGCTGCCCAACAAA
 AA
 AA

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Signal peptide:

N-glycosylation site.

Glycosaminoglycan attachment site.

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-21	32-38	42-48	50-56	60-66	61-67	64-70	74-80
90-96	96-102	130-136	140-146	149-155	152-158	155-161	
159-165	163-169	178-184	190-196	194-200	193-205	218-224	
236-242	238-244	239-245	240-246	245-251	246-252	249-252	
253-259	256-262	266-272	270-276	271-277	275-281	279-285	
283-289	281-290	287-293	288-294	291-297	292-298	295-301	
298-304	301-307	311-317	312-318	313-319	322-328	323-329	
325-331	343-349	354-360	356-362	374-380	381-387	383-389	
387-393	389-395	395-401					

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCGACAGAAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCAGTGTTTCC
CAGAGCCCCAAAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTCATCGTTTTATGCCACCTGACACCATCCGGTCTATCACCAGTCCCTCAGCTGCCA
TTGCACCAAGGATAATCTCTTCATCAGGTTCTTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGCCGCTGACAAGTGGAGCGCCACCGTCGGATGCTGACGCCGCCCTTCCATTTCACATCCT
GAAGTCTTATATAACGATCTTCAACAAGAGTGCAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGGACGAGTCGCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGGGCCAGTGAATATATTGCCACCAT
CTTGAGAGCTCAGTGCCTTGTAGAGAAAAGAGCCAGCATATCCTCCAGCACATGAGATTTCTGT
ATTAACCTCTCCCATGACGGGCGCGCTTCCACAGGGCTCGCCGCTGGTGATGACTTCACAGAC
GCTGTCAATCCGGGAGCGCGCTCGCACCTCCCACTCAGGGTATTGATGATTTTTCAAGACAA
AGCCAAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGAGAGCCATGACACCAG
GCCAGTGGCTCTCCTGGTCTGTACAACTTGGAGGCAACCAATACCGAGGCGCTGCGG
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCAATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCAGTCACTCCCAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACACCCAACTGTGTGGCCGATCCTGAGGTCTACGACC
CCTTCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACTCTGGCTTTTATTCCTTCTCCGCA
GGGCCACGAACTGCATCGGCGAGCGTTTCGCTATGGCGGAGATGAAAGTGGCTCTGGCGTTGAT
GCTGTGCACTTCGCGTTCTGCCAGACACACTGAGCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCACTGCACTTTCAGC
CCATCCACCTGTTTTTTTGAGATTGTCTGAATAAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVGSGWLLARILAWTYAFYNNCRRLQCFFQPPKRNWFWGHLG
LITPTEEGLKSTQMSATYSQGFTVWLGPIIPIFVLCHPDTIRSIITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFRAC
RLVHDFTDVIRERRRTLPTQSIDDFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAFFISRCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVIALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACCTTACCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCAITGACGTGG
TACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTTTTGCCGTGTGCTTGC
ATAATTTCATGGCCAGTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAACATATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAACAATGTGTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATCCCATGAATACAAACCTATTGAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 57

CGGCTCGAGCTCGAGCGGAATCGGCTCGAGGGGCACTGGAGCACCCAGCAGGCGCGCAACATGCTCTGTCTGTGCGCTG
TAGCTGCGGCTCATCGGGAGACCCAGCCAGGATTCAGTACTTTGAGTCGAGAGGGGCTCCCTGCAGAGCTGAAGTCC
ATTTTCAGGCTCAGTGTCTTCATCCCTCCAGGAATTTCCACCTACCCGATCGAGCAGAGAAATTTACAGACT
GGAGATAAGGACTTGATGGCGAGCTAGACTTTGAAGATTTGTGCATATATCTCGAGATCATGAGAGAGAGCTGAGG
CTGGTGTGTTAAGATTTGGACAAAAGATGATGGAGCATTTGACGCGCAGAGGATCATGAGTCCCTGGCGAGACTGG
GGAGTCAGAGTATCTGACAGCGGCGAGAAATTTCCAGAGGCTATGGTAAAAAGCGCACGTGACCTCATCTGG
AAGAGTGGAGAGACTACCACTCCTCCACCCCTGGAAACATCCCGAGATCATCTCTACTGGAAGCATCCACG
ATCTTTGATGTGGTGAGAACTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGACAGCGGGATGTGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCCGATTCAGAGACTGCAGGCCCCCTGGACAGGCTCAAGTGTCTCAT
CAGGTCCATGGCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTCGAGAGGAGGGGCCAG
TCATCTCTGGCGGGCAATGGCATCAAGCTCCTCAAAATTTGCCCGGAATCAGCCATCAATTCATGGCTATGAGCAG
ATCAAGGCGCTTGTGGTGTGACCCAGGAGACTCTGAGGATTCACAGAGGCTTGTGGCAGGCTCTTGGCAGGGGCC
ATCCCGCAGAGCAGATCACCCATGGAGTCTGAAGCCCGATGGGCTCGGAGACAGGCGCATCTCAGGA
ATGCTGAGACTGCCCAGGAGGATCTTGGCAGAGAGGGGTGGCCGCTTCTCAAGAGGCTATGTCCCAAGATGCTG
GGCATCATCCCTATGCCGCGATCGACTTGCAGTCTACGAGACGCTCAAGATGCTGGCTCGACACTATGCACTG
AACAGCGCGACCCGGCGGTGTGTGCTCTGGCTGTGGCACCATGTCCAATGCTGTGGCAGCTGGCCAGCTAC
CCCTCGGCCCTAGTCAGGACCGGATGAGGCGCAAGCTCTATTGAGGGCTCGCGAGGTGACCATGAGCAGCCTC
TTCAACATATCTCGGAGCCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCACTTCATGAAGTCTCATCCCA
GCTGTGAGCATCAGCTAGCTGGTACGAGACCTGAAGTACCTGGGCTGGCTGAGTGGCTGGAGCTGGAGGGGAGGCG
GGCCCGAGCATGGAGTCTGATCTGGGCGGAGCTGGGCTGGAGCAATCTCATCTGTGAATGTGCACAT
AAGCTGTCTGAGCCAGCTGTGAACCTTAGACGCAAGCCGAGGAGGGTGGGAGAGCTGGAGCCAGGCTCT
GTCTGCTGACCCAGCAGACCTCTGTGTGGTTCAGCGAAGCAACAGCAATTCAGGCTCGAGGCTCGAGGCTG
CTCGGGCTCAGTGTGAAGCAGGACATTTCTGAGTGCCTGCCAATGTAGCTTGGAGCTGGAGCGCCGCT
TAGTGTCTCATTTACCTTGCAGCAGCTGTGGCAGCGGCCCTGCCCTGTGGTCTGCCGTCACTCGCTGTGC
CCCTGTGCTGCTGCTGTCTGCTGAGGTAAAGTGGGAGGAGGCTACAGCCCAATCCACCCCTCTGCTCAATCCC
ATATCCATGTGAAGGTGAAGTGAAGTGGAGTGGCTCGAGGCTCGAGGCTTCCCAAGCTCAGGCTTACGCGCAATG
TGTGAAGAGAGGAGAGGATCTGGCTGTGGTCACTGGAGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCT
CTTGGAGATGAGGGGGCTGGCTGGCTGGCTGGCTGCACAGAGAGCAAGTGTGGGCTCATGGTGTCTGAGCT
GGCTGGAGGCTGTGAGATGGGCCCATCTCAGAACCACTCACTGTCCCACTGTGGCATGAGGCGATGGAGCA
CCATGTTTGGGGCAGAGGCGAGAGCTTTGTGTCTGGGGAGGAGGAGAAAGGTGTTGGAGGCTTAATTAATGG
ACTGTGGGAGAAAGGTTTTGTCCAGAGGACAGCCGACCAATGAGCGACTTCTGTGCTCAGAGGAGACAGAGG
GAGCAGAGCTTGGCTGACTGCTCAGAGTCTGTCTGACGCCCTGGGGGTTCTGTCCACCCCGCAGGGGCGAGC
GGGACCAAGCCCACTTCCACTTGTGTGACTGCTGGAGCTAATTTATTTGTATTTATTTGAACAGAGATGTCTCT
AATCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
GATTTACCTTCCCAAGCCGCCATGTGGATGGGAGGAGAGAGAGGGGGCTTGGCCGCTGCTCATCTCATCT
GTCCAGAGAAATTCCTTTTGGGACTGGAGCAGAAAGGCGCCAGAAAGGACGACCTGGCTCTTCTTCTTGGCAG
GTTGGGAGAGGCTTGGCCCGAGCCTTAGGATTCAGGGTTGACTGGGGGCTGGAGAGAGAGGAGGAACCTCAAT
AAGCTTGAAGTGAATCCAGTATTTCTCGGCTGCGAGGTTTCTTATTTCACTCTTTCTGAATGTCAAGCGAG
TGGGTTGCTCTCACTGTGAATTTGTGGTGGCGGGGGCTGGAGGAGGGGTGGGGGGCTGGCTCGGCTCCCGAGC
CTTCTGCTGGCTGTGTGAATGGCGGCACTGGAGCTCAAGGTGTGACCTTCATTCGACAGATGACCTGA
TGAGGAATTCATAGAGTCAAGATCAATGCAAAATTTGTTATATATATATATATATATATATATATATATATATAT
CAAAATAGAAAGAAATGGAGCTTAGAGTGTGATTTAAAGCAGGCTTCAATAAAGTTGTTTCAAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEATQTEFQYTESKGLPAELKSIKLSVFIQSQEFSTYRQWKQKIVQAGDKDLG
QLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQAEKILKSMKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTEVERQTGMWHRHLVAGGG
AGAVSRTCTAPLDRKVLIMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQDETIRIHERLVAGSLAGATAQSSIYPMEVLKTRMALRKTGOYSGMDCARR
ILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGSCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCTTGGGGCAGATCCTCTTCTGGAGCATAAITAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCAGAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAACTCTGAGCTGCACCTTTGAACCTGACATCAAACTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTAGGCTTGGTCCATGAGTTCAAAGGAAGCAAGATG
AGCTGTGGAGCAGGATGAAATGTTAGAGGCGGACAGCAGTGTGTTGCTGATCAAGTGTATAGTT
GGCAATGCTCTTTGCGGCTGAAAAAGTGCACCTCAGAGATGCTGGCACCTACAAATGTTATAT
CATCACTTTCAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAAGATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGGGTGTGAGGCTCCCGATGGTTCCTCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTGTACAGGGAGCCAACTTCTCGGAAGTGTCCAATAC
CAGCTTTGAGCTGAACCTCTGAGAATGTGACCATGAAGTTGTCTGTGCTTACAATGTACGA
TCAACAACACATACTCTGTATGATTGAAATGACATTGCCAAGCAACAGGGATATCAAAGTG
ACAGATCGGAGATCAAAAGCGGAGTCACCTACAGCTGCTAACTCAAAGGCTTCTGTGTGT
CTCTTCTTTCTTGGCATCAGCTGGGCACTTCTGCTCTCAGCCCTTACCTGATGCTAAATTAAT
GTGCTTGGCCACAAAAAGCATGCAAGTCAATGTTACAAACAGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTGTATATTTCTGGAGGAATGAATTATATCTAGAAGTCTGGAGTG
AGCAACACAGAGCAAGAAACAAAAAGAGCCAAAGCAGAGGCTCCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAAATAATCATGTGAACATAGACAAGTGTGTAAAG
GTGATAAGTAAATGCAGCTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTATATGTGC
TGTAAATGTGCTCTGAGGAAGCCCTGGAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCTTAAGACGCTGCTAATGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGSTCAAATGATTCACCTTTTATGATGCTTCCAAAGTGCCTTGGCTTC
TCTTCCCACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTAGACACCTTCTTTTAAACAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEPKGKDELSEQDEMFRGRTAVFADQVIVGNASLRKKNVQLTDAGTYKC
YIITSKKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSMIENDIAKATGDIKVTESIKRRLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGT CAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
 CACCAGGAGCCCTCCGGGTAGCTACTACCTGGACCCCAATAGTGGAGGGCAGTATGTGTAGT
 GGGCTACCCCTGGTGGTGGTTATGGGGGTCTGCCCTCGGAGGGCTTATGGACCACCACTGG
 TGGAGGGCCCTATGGACACCCCAATCTGGGATGTTCCCTCTGGAACTCCAGGAGGACCATATG
 GCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCTTACGGTCCCAGCAGCCT
 GGGCTTTATGGACAGGGTGGGCCCCCTCCCAATGTGGATCCTGAGGCCCTACTCTGGTTCCAGTC
 GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
 ATTGGTCTTCATTCAATGATGAGACCTGCCCTCATGATGATAAATCATGTTGACAAGACCAAGTCA
 GGCCGCATCGATGCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAATCTCTT
 CCAGCAGTATGACCGGACCCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
 AAATGGGCTACAACTGAGCCCCAGTTCACCCAGCTCTGGTCTCCCGCTACTGCCACGCTCT
 GCCAATCTCGCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTCTGCACAGA
 GGCTTCCGGGAGAAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
 CCATGACAGCTTCTCGATGCTATGACCCAACCATCTGTGGAGGTGGAGTGCACAGGGACCTT
 TCCTGGCTCTTAGAGTGAGAGAGTATGTGGACATCTCTTCTTTCTCTGCCCTCTAGAAGAAC
 ATTCTCCCTTGCTTGATGCAACACTGTTCAAAAGAGGGTGGAGAGTCTGCATCATGCCACCA
 AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCCCTGATGGAGGAGAGGATAGAAGTTGA
 ATGTCTGATGGCCATGAGCAGTTGAGTGGCAGAGCCTGGCACCAGGACAGGTCTTGTAAATGG
 AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAAGTTCATCGGCTGTTACC
 GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAAT
 CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTTGGATTCTGCCACACCATAAAT
 CTTGTGTGTAACTTCTAGTGCCTGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
 CTTTGGCCAGGCTCTGCCCTCTCAGCTGGGACCCCTCACTTGCTGCCATGCTGCTGCTCGGCT
 TCAGTCTCCAGGAGCAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCAATTTTTTTC
 ATTTGGGCCAAAAGTCCAGTGAATGTAACTTCAATAAAGGATGAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPAGGGPYG
HPNPGMFFSGTPGGPYGGAAPGGPYGQPPPSYGAQQGLYGQGGAFPNVDPEAYSWFQSVDSDH
SGYISMKELQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFIIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCGCGGTGGCAGGGAGCTGCGCTCCTCTGGGCTGCTCCTGGTCTGCTTCATC
 TCCAGGCCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCAAAACCTCGGGACC
 AACTTGCCCTCAGCTCGGACAACCTTCTCCACTGGCCCCCTCTAACTCTGAACATCCGAGCCCCG
 TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCTCCATCAGATG
 GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
 GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCGGGCAGTGGCCCTT
 TGCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACAGGAC
 TCGGAGTCCAGACGACTGCCCCGTTCTAATTCACTGGGAGCCGGGGAAAAATCCTTTCCCAAAG
 CCTCCCTGGTCTCTCATCCACAGGGTCTGCTGATCACCCCTGGGGTACCTGAATCCAGTG
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGAACGAGGCCCATGCCACACCTGAGGGA
 ATCTGGGGTATCAATAATCAACCCCGAGGTACAGCTGGGGAAATATTAATCGGTATCCAGGAG
 CAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGAAATATTAATCGGTATCCAGGAG
 GCAGCTGGGGAAATATTCATCTATACCCAGGTATCAATAACCCATTTCCTCCTGGAGTCTCCGC
 CCTCCTGGCTTCTTCTTGGAAATCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
 GGGCTAGACGACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACATATCCCACTTTTCAGTGCTCCCC
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 64

MQGRVAGSCAPLG LLLVCLHPLGLFARSIGVVEKVSQNFNTLPQLGQPSSTGSPNSEHPQPAL
DPRNDLARVPLKLSVPPSDGFFPAGGS AVQRWPPSWGLPAMDSWPFEDPWQMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSFEASLLHQDSESRRLPRNSL GAGGKILSQRP
PWSLIHRVLPDHPWGT LNFVSWSGGGGPGTGWGTRFMHPHEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNFPPGVLRRPPGSSWNIPAGFNPFPSPRLQNG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCTCCATCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCTTTGTCTCCTCTTGACCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
GGGCTTTGGGCCGGGGATGCAGGAGGCAGGCCCGACCTGTCTTTCAGCAGGCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSLPLVLLLTLLGSSHGTPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHVVVNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTCGCAGGTGCCCTTGGCCGGAGATCGGGTAGGAGGGGCGAGCGCGAGAAAGCCCC
TTCTTCGGCGCTGCCAACCCGCCACCCAGCCCCATGGCGGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTCTGCTGGCCGCTGGGGCCGAGCCTGGGGGCAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCCTTCATCCACAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCTTGCTCCTGGCTGTGGGCTG
GCACGTGTGGTGGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTCAGGGCT
GCCTGCCCATCTAGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCTCTCTGGGCAGTCAGATCCACCCAGTCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTTATATATTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLFLLARWGRAWGQIQTTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKILREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATACTAGAGAGCAAAATGGGGTTATTGAGAGGTTTGTGTTTCTCTTAGTCTGTGCTGTGTCACCAAG
 TCAATACTCTCTTCATTAGCTGATTAATTAATGGCTTGAAGATATTGTCATTGTTATAGATCTGATGTGTGCGAGAA
 GATGAAAAAATAATGAACAANTAGAGGATATGGTGAATCTACAGCTTCTACGTACCTTTGAGGCCACAGAAAAAGA
 TTTTTTTCAAAATGTATCTATATTAAATTCCTGAGAAATGGAGGAAAAATCCTCAGTACAAAGGCCAAAAATGAA
 AACCTAAAGATGCTGATGTATATAGTTGCACCACTACACTCCGAGTAGAGATGAACCAACACCAAGCAGTTCACA
 GANTGTGGAGAAAGGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAATAAGATATGGACCA
 CGAGCAAACTGTTTGTCCAGAGTGGGTCACCTCCGCTGGGAGTGTGATGATGATACATGAGATCAGCTTTC
 TACCCTGTCTAGTCAAAAAAATCGAAGCAACAGGTGTTCCGAGGTATCTCTGGTAGAATAGAGTTTATAAGTTCT
 CAAGGAGGCGCTGCTTAGTAGAGCACTGAGAAATGATTCTACACAAAACCTGTATGGAAAAATTGTCAATTCTTT
 CTTGATAAAGTACAAACAGAAAAAGCATCATAATGTTTATGCAAGTATTGATTCTGTGTTGAATTTGTAAACGAA
 AAAACCCATATACAGAGTCTCCAGCCTACAAAACATAAAGTGAATTTAGAAATACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACTCTCCACCTGTCTCTCAITGCTGGAAGTCAGTCAA
 AGAATTTGTGTCTTATTCTGTATAGCTTGGAGCATGGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGGAGCTGTGAAAAGTGGTCTGGTGGGATGGTTCACCTTGATATGATCTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAGACAGTGATGAAGAAACACACTCATGGCAGGATACCTACATATCTCTGGGAGGA
 ACTTCCATCTGCTGGAAATAATATGCAATTCAGGTGATTGGAGGCTACATTTCCCACTCGATGGATCGGAAGTA
 CTGCTGCTGACTGTGGGAGGATAACATCGCAAGTTCTTGTATTGATGAAGTAAACAAATGGGGCCATTGTTCTAT
 TTTATTTCTTTGGGAGAGCTGCTGATGAAGCAATAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCGAGTGAAGCTCAGAACATGGCTCATTTGATGCTTTTGGGGCTTTACHTCAGGAATACTGATCTCTCCAGAGG
 TCCCTTCAGCTCGAAGTAAAGGATTAACCTGAATAGTAATGCTGGTGAACGACACTGTGTAATTTGATGATACA
 GTGGGAGAGGACAGTTCTTCTCATACATGGAGCTGTGCTCCCACTGATTTCTCTTGGGATCCAGTGAAGA
 ATAAAGGAAATTTACAGTGGATGCACTTCCAAATGGCTATCTCAGTATTCCAGGAACTCGAAAGTGGGCACT
 TGGGCTACAACTTTCAAGCCAAAGCGAACCCAGAAACATTAATATAGTAACTCTCGAGCAGCAAAATTTCTCT
 TGGCCCTCAATCAGATGAATGCTAAATGAATAAGGACGTAAACAGTTTCCCAAGCCCAATGATTGTTACGAGAA
 ATTCTACAGGATATGTACTGTTCTTGAGCCAAATGTGACTGCTTTCAITGAATCAGAAATGGACATACAGAGGTT
 TTGAACTTTTGGTAATGGTGGAGGCTGATTTCTTCAAGATGATGGAGTCACTCCAGGTATTTACAGCATAT
 ACAGAAATGGCAGATATAGTTTAAAGTCCGGCTCATGGGAGGAAACACTCCGAGCAATAATCCGGCTCCA
 CTGATAGAGCCGCTACATACCGAGCTGGGTAGTCAACGGGAAATTTAGAGCAACCCCGAGACCTGAAATGAT
 GAGGATACTCAGACCACCTTGGAGATTTCAGCGAACGATCCGAGGTGCAATTTGTGATACAGAGTCCCAAGC
 CTTCCTTGGCTGACCAATCCCAAGTCAAAATCAGAGCTTGAAGCCACATATTTGCTGATTAAGATATGATTT
 ACATGGACAGCACCAGGAGATAATTTGATGTTGAAAAAGTCAAGCTTATATCATTAAGATAAGTGAAGTATTCTT
 GATCTAAGAGCAGTTTTGATGATGCTCTCAAGTAAATACTACTGATCTGTCAACCAAGGAGGCCAATCCAGGAA
 AGCTTTGCTATTAAAGCAGAAATACTCAGAGAAATGCAACCAATATTTATGCTATTAAAGTATAGATAAA
 AGCAATTTGATATCAAAATATCCAAAGTCAAGATATTTGTTATCCCTCAGGCAATCTGATGATGATG
 CCTACCACTACTCTACTCTACTCTCTGATAAAGTCATAATCTGGAGTAAATTTTCCAGCTGTGATTC
 TGTGTGATTGGGTCTGTTGTAATGTTAACTTTATTTAAAGTACACCACTTGACCTTAACGAGAAAAAATCTCT
 AATGAGACTAGAGAGGTTTTAAAAACAAACAAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT
 GTGTGATCATAACTCTAAATAATTTTAAAGTTCGGAAAAAGATACCTTGATTAATAAAAAACCTCATGGATA
 TGTAAATCTGTCAAGATTAATAATTAATAGTTTCATTTATTTGTTATTTTATGTTAAGAAATAGTGATGAGAAAG
 ATCTTTTATCATGATGATACCTGGTGTATATTTGATGATGATGATTTTGGAAATGATTTTCAATTTGATCAAG
 GAATTAATAATCATCTATCTGAGTAGCAAAATACAGTAAAGGAGAGAAATAACACATTTTGGAAAAAAGAA
 AA

FIGURE 70

MGLFRGFFVLLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFE
ATEKRFFFKVSIILIPENWKENFQYKRPKHENHKGHADVI VAPPTLPGRDEPYTKQFTSCGEKGEY
IHFTPDLLLGGKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQFFYRAKSKKIEATRCSSAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFFDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDLNR
MNQAAKHFLLTQTVENGSWVMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLTGDEDNTASSCIDEVKQSGAIVHFIALGRADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAMMNDTVIIDSTVGKDTFFL
ITWNSLPFISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNQGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQTILEDFTSRTASGGAFFVSVQVPSLPLPDQYPPSQITDLDTVHEDKIIILTWAPGDN
FDVGKQRYIIRISASIIDLDLDSFDDALQVNTTDLSPKEANSKESPAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIQAQVTLFIPQANPDDIDPTPTPTPTPKDSHNSGVNISTVLVSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCTTATGTTGAAACCCCTGGGAGTAGTACTGACAGCAAGACCGGGAAGACCATACGTCCTCCCGGGCAGGGGTGA
 CACAGGGTGTCTATCTTTTGTATCTCGTGTGTGGCTGCTCTCTTATCTCAAGGAAAGACCGCAAGTATTTTGGACCA
 GAGGACAGATGATGTACGCCACCTCTCAACCTCCCTCTTGAAACCCAGTATTTGCCAGATTTACTAGAGATGTGCA
 ACTCAACACGCAAGCGGCTCTCCGGCTTAACTTGTGTGTGGAGGAGAACCTTGTGTGGGGCTGCTTCTCTTAGCA
 GTGCTCAAGAGTGACTTGCTGTGGGGGTGACCAAGAAAGGAAGTCCCTCTTCTGTGTGGCTGACATCAGGAA
 GCTCTGTATGGAGTAAAGTGAAGCTTGTGAGATTCTACCTCATGATTTGCTCTGTGCTGCAAGATCATCTTAA
 ASTAGAGAGCTGCTCTGTGTGTGTAACTCCAAGAGCGAAACCTCTCTAAGAGAAATGGATCAAGCAACTC
 CGGGGGCCCCCAACCGCATGCTTCTGTGTGTGCTAGCCAGGGAAGCCCTTCCGTGGGGGCCCGGCTTTGAGGGATGCC
 ACCGGTCTGTGAGCGCATGGCTGATTTCTGAGTGTGATGTTTGGCGGGGGCTGCTGCTGTGGNTTTCCCGGGTGTG
 GTTTTCTGTGTGCTCTGCTGTGTGTATCTCTGTCTGTATCTGCTGTGCTGCTGACCCCAAGGTGACAGGAGAG
 CTGGCAGCTGCCAGGCGCAACGCCACCGGGGAAGAGGGGTACAGGCCCTCTTCAAGGATGGAGAGCAACAC
 CGCACTACTGTGAGCAGCTGGAAGCGCAGATGCTACAGCTCAAGAGGAGGCTGCAAGAGAGAGTGAACACTCAG
 AATGGGACATACAGGCGCAGCATGCTGT
 CTGGCTCTCTGCTGCTGTGAGAGGTGCAAGGCGCAGAGGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 CTTTTCAGATGCTTACTCTACMAAGGTGTACAGCTGGAGACTGGCTTACCGCCACCCCGAGGAGAGGCTGTG
 AGGAAGGACAGCGGATGATGT
 CCCAATCAGCTCTCTTACAGGGCTCTGTATTTATAGAAGGGATCTACCGAACAAGAGGACAAAGGGACATTGT
 GAGCTCACTCTCAAGGGGACCAACAAACAGATTTCAACCGCTCATCTTATTTGACCATTCAGCCCATCATGA
 GTGAAATGAAAGCTCAACATGGCACAAGGCTTATCAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 CGGAGTCTGCAAGANTTTCAAGGAGGTGCATTTGAGGAGGTGAGAGGTGCACTCTCTGTGTGTGTGTGTGTGT
 AAGAGAGAAATTAATGAAGTCAAGGATACTTGAAGCACTTCCAAAGTGTGCACTTCAGGAACTTACCTTCACT
 CAGCTGAATGGAGATTTTCTCGGGAAGGAGACTGTATTTGGAGCCCTCTGTGAAGGAGAGCACTGCTCTCTC
 TTTTTCTGT
 GTATTTTATCCAGTCTTTTCACTGATCAATCTGTGGCATAATATACGGCCACATGATGCACTCCCTCCCTTGTGA
 CAGCAGCTGTGCATAGAAGGAACTGGATTTGTGAGAGACTTTGTATTTGGATGAGTGTGCAATTCGTGTGAGC
 TTCTCAATTAAGTGGT
 CAGCAACTCTCATAGGTGCAAGGAGGCTGT
 CTGACCCCGCAGCATACAGATGTGATGCACTGCAAGGCGCATGAACGAGCATCCACGGCAGCTGGGCACTGCT
 GTTGTGAGGACAGATAGAAGGCTCACTTGCNAACAGAACGNAAGCAATGTGCAAAAAGCACTTCCCAAG
 GAGGATGTGGGAGACATTTTCTTCTTGT
 GGACGACAAAAGATTTGAGT
 AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCGAGTTGCACTGTGAAGTGTGCAAAAGGCAAGATGCTGTGT
 AGATTAATAGCTAATGT
 AATATTCATGATTTTGTGAGGCTTTGTAAATATCATATGACATAGAGGCAAGATTTTGTGTGTGTGTGTGT
 GCTATCAGCAGGCTCTGATTTCTGAATATCTAATATCAAGAGGCGAGAGGAGATGAGCTTATTAATGATAT
 AGTGAATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TAACCAAAATTAATCTGCTTATCTTTTGT
 TTTTTTCCCTGTGAGTATAGTCTGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 ATTTTATATTTTAAAGAGATCTTTGAGATGATATGAGACTTTTCAATCAAGGATCAATTTGATGCTATTT
 CCAGGAGCTGCAAGTGTGATTTCTGTGAGGACTGATGTGAGGATTTGAGGAGGAGGAGGAGGAGGAGGAGG
 ATACAGAGCTACAGATCTTCTGAGATGATTTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
 ACTTTCTGCTTACAGAAAGGAACTCATCAGCTGCTGATATCTGATGATCACTAAAGATCAGAAACCAATTT
 CTCTCAGAGATGAGGAGGCTTTTCTTACTGTGTAAATTAACCAAGTATACGCTGTGAACCAACAACTCTTTTC
 AAACAGGGTGTCTCTCTGCTTCTGGCTTCATTAAGAAAGATGAGAAATATATATATATATATATATATAT
 GAAGATCAATCCATCTGCCAGATCTAGTGGATGGAATTTTGTGATGTATGATCCACCCAGCGAGGTGGAG
 TACTGATATTTTAAATTAAGAGTCTTACTCATCAGCAAGATGCTTCAAGAAATGATTTTATTAACATTT
 CAAGTATTTTAAATTAAGAGTCTTACTCATCAGCAAGATGCTTCAAGAAATGATTTTATTAACATTT
 ATGCTGAGCTTATTAATCTTTGAGATGCTTCTGT
 ATTCAAGCTGTGTGTGTGTAAATTAAGATGATTTGATTTGTATGATGATTAATTAATTAACACAGG
 CCTAAGTGAAGTGTGTGTGACAGCTAATTAATATGATTTTGTGATGATGA

FIGURE 72

MMVRRGLLAWISRVVVLVLCCAISVLYMLACTPKGDEEQLALFRANSPTGKEGYQAVLQEWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGVQASDAAGLGLDRSPFEKTQADLLAFLHSQVDRAEVNAGVKLATEYAAVFPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNPAENSPNHRPYTASDFIEGIYRTERDKGTLVELTFKGDHKEHFKRLILFRFPSPIMKVKNELNMANTLINIVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEINEVKGILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTRCLNTQPGKKVFPVLFESQYNPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYRKYLHSLIVVRTPVVRLFHLWHEKRCMDELTPQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

APP ID=10063713

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNHGTVQLSLPSTLYLGGLPKRYVAAQLHLHWGQKGSPPGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHLSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPFCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLNKRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGTGCTGCTGTGTGCTCCTGGCGCGCCTTGGGGACGGGAGTTCCCTGTGTCT
TCTGGTGGTTTGCCTAACACTGCAAAATCACCCTCTTATCCATCAACATGAAGAAATGCTCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATCACAA
ATTGGCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTGTGCTGACAGCTCC
AGAGAAGTGAAGAGAAATCCAGAAGACCTTCTGTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTGAATACTAAATCAAACAGAACGTGGTCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCAGCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCTTCTGTCOC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAANGATCAATCAT
CAGAGTCAAGGCTAAATCATCTTCTGGTAIGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTGGCAAAGAGAAACCCAGCAAATTT
GATTTTGATTTATGAAATGAATTTGACAAAAGATTCTTTGTGCTGTGAAATAATCGTGATTA
ACTTTATCACCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCAGCGGGAACTGAGGCCCTCAGGA
GGAAGAGGAGGTGAACATTAGGGTATGCTTCGCATTGATGGAATTTTGTGACTCTGAAG
AAAACAGGAAGTACTTCTCTACCAGCAAGAGTCCCTCAGCAGCAACATACCCCGGATATA
ACAGTCATTGAATATGAATATGATGTGAGAACCAGTACATTTGTGCGGGGCTGAAGAGCAGGA
GCTCAGTTTGAGGAGGAGGTGTCCACACAAGAACATTATTGGAGTGCAGGCAGCGTTGGCAG
TCTTGGGCCCGCAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGOG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTGCAGTGGGA
TCCCCAACTGGCAGGCTGTGATTCTCTCGCTGTCCAGCTTCGACACAGATTACAGAGGCTGCG
AGCCTTCTGAGGGGATGGGCTCGGAGAGGAGGTCTTCTATCTAGACTCTATGAGGAGCGGCT
CCAGACAGGCCACAGGAGAAATGAACTTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTCAGAGTGAAGAAATGATGCCAACACTTCCCTTTTGCCCTTTTGTTCCTGTGCAACAGTIGAG
TCACCCCTTTGATCCAGCATAAAGTACCTGGGATGAAGAGTATTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTTGGTTCATGCAITGA
GGTCTCTTAACATGATGGTGGGCTCTGGAGTCCAGGGCTGGCGGTTGTCTATGACAGAGAA
AGCAGTCAATAAATGTTTCCAGACTGGGTGACAGATTATTCAGGTGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLFVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAPSEKQCARTLK
DQSSEFKAKII FWYVLPISITVFLFSVMGYSIYRIHVHGKEKHPANLLIYGNEFDKRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEDVRTTDCAGPEEQELSLQEEVSTQGTLLSQ
ALAVLGPQTLQYSYTPQLQDLPLAQEHDTSEEGPEEPSTTLVDNDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEAPADRPPEGENETYIMQFMEENGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTTCTGGCATCCTGCATCTGCTGCCCTCTGACAC
CTGGGAAGATGCGCGGCCCTGGAGCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCTCAGTCCCAGTGCAGTTCTCATCTCGGCCAAAAGTCATCAAGGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAAGCCACAGCATCCTGCAGCAGCTGCCGTGCTCAGTGCCATGC
GGGAAAAGCCAGCGGAGGCATCCCTGTGCTGGGAGCCTGGTGAACACCGTCTGAAAGCACATC
ATCTGGCTGAAGGTATCACAGCTAACATCTCCAGCTGCAGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGCTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTGAAGTGGC
CCCACCCGCCCTGTCCTCAGTGAAGTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA
TAAGCTCTCCTTCTGTTGAAGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAGAACAGCTGTGTCCGTGATCGAGGCTTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTCAGCTCTACCTGGGGGCCAAGTGTGTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTTCCCTGACAAATGCCACCTGGACCAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTCTCTCCAGAAGA
ATTATGTCCTCTGTTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTCTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTCAACCTGGGATCGAAGCCAGCTCGGAAGTGT
TTGTGGAACCAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGATCCAGCTG
ATGAACCTTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCTGCTGCCGAACAGAATGGCAAATTAAGATCTGGGTCCCAAGTGTGATGTTGGTGAAG
CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCTCC
TTGTGGAACCAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGATCCAGCTG
GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGFWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSI LQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTEAQATIRMDTSASGPTRIVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLDSQGKVT
KWFNNSAASLTMTLDNIPFSLIVSQDVVKAAVA VLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTP EFFIDQGHAKVAQLIVLEVFPSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFPDVLKNIITEIIHSILLPNQNGKLRS GVPVSLVKAIG
FEAAESSLTKDALVLT PASLWKPFSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGACCAAGAGC
TTCAGCCTGAAAGACAGGGAGCAGTCCCTGAAAGCGCTTCTACTGAGAGGTCGCCCATGGCCTCT
CTTGCCCTCCCAACTTGTGGGCTACATCCTAGGCCCTTCTGGGGCTTTTGGGCACACTGGTGCCAT
GCTGCTCCCCAGCTGGAACAAGTCTTATGTGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCGCTGACATC CAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTC
ATTCTGTGGCTGGAATCTTTCATGGGATCC TACGGGACTTCTACTCACCACCTGGTGCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTAGGAACACAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAGGCCAAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTCC
TCACCTTGCTGCTCCCCGCCCCTAAGTCCCAACCCCTCAACTTGAAACCCCATTCCTTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCCCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCCTGACTGACCTCTGTGATCAAGACCCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAAGTGGCTTTTGTGGGCATTGCTCAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTGTTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACCTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAIISSLACIIISVGMRICTVFQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNLHGILRDFYISPLVPDSMKFEIGEALYLGIISSLSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGGTCCGGCCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCATCTCTCCCTCCTTTTC
 CCGCGGTTCCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCTGCCCCCTCTTCTTC
 GCCACCGCTGCTTCTTGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGTATCTGTGGCCCTGTGCTCCGTCTCTTTCTGTCTCCCTTCTCCCGACTCCGCTCCCGG
 ACCAGCGGCTGACCTGGGAAAGGATGCTTCCCGAGGTGAGGTCTCTCTCTCTTGTCTGGGA
 CTGCGCTGCTCTGGTTCCCTTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCTTTT
 CCATGGGAAGAGATACTCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCTGATGT
 ACTGCTTGCCTGTACTGCTCAGAGGGCGCCCATGTGAGTGTGTACCGCTTCCACTGTCCGCT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAGTCTTGCCAGCACACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTCCCTCCCGCTGCCCAACCAAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCTCACAACTGCCCCGAACCAAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCTCTGCTGCCAAGCTGCAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCATGGGGTGAGACATCTCAGGATCCATGTTCCAGTGATGCTGGGGAAG
 AGAGGCCCGGGCACCCAGCCCCACTGGCTCAGCGCCCTCTGAGCTTATCCCTCGCCACTT
 CAGACCCAAAGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCT
 GTGTGCATGGCGGGAAGACTACTCCACGGGAGGTGTGGCACCCGGCTTCCGTGCTTCCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACTGTCC
 CACCGAGTACCCCTGCCGTCACCCGAGAAAGTGCTGGGAAGTGTGCAAGATTGCCAGAGG
 ACAAGACAGACCTTGCCACAGTGAGATCAGTTCTACAGGTGTCCCAAGGCACCGGCCGGCTC
 CTGCTCCACATCGGTATCCCCAAGCCAGACAACTGGCTCGCTTTGCCCTGGAACACGAGGC
 CTGGACTTGGTGAGATCTACCTCTGGAAGCTGTTAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACAGCCAGAACTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCGACTGTCTGCTGGCCCCACGAAGTCACT
 GGAACGTCTTCTAGCCAGACCCCTGGAGCTGAAGTCAAGGCCAGTCCAGACAAAGTACCAAG
 ACATAACAAGACTTAAACAGTTGAGATATGAGCTGTATAATTGTTGTTATTATATATAATAAA
 TAAGAAGTGCATTACCTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSLLGLALLWFPLDSHARARPMFCLFHGKRYSFGESWHPYLEPQGLMYCLACTCSE
GAHVSCYRLHCPPVHCPQFVTEPQQCCPKCVPHTPSGLRAPPKSCQHNGTMYQHGEIIPSAHELF
PSRLPNQCVLCSCTEGQIICGLTTCPEPGCPAPLP,PDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPIRHFPRKAGAGSTTVKIVLKEHHKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRTCTEYPCRHPKAVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPDNLRFALEHSEADLVEIYLKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAECHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTCAGAACAGCGAGTTTGGCCCTCCGCTCACGCAGAGCCTCTCC
 GTGGCTCCGCACTTTGAGCATTAGGCGAGTTCTCTCTCTCTCTAATCCATCCGTCACTCTCTCTGTCA
 TCCGTTTCCATGCCGTGAGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGTGTITGGGCGAGACAAGCCTGTCCAGGCCTTGGTGGGGGG
 GAGCGAGCATTCTCTGTTTCTCTGTCCTTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAAGCATTATGTCAGATGCCAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGAAACATTACT
 GTGTTGAGTGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGCTTTTACTACCAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCACTCTCGGGCTGGTTCCCGGGCCACAGCGAAGTGGAAAGTCCACAAGGACAGGATTTGTCC
 ACAGTCTCAGGACNAACAGAGACATGCATGGCTGTTTGTATGTGGAGATCTCTCTGAGCCTCCAGAGGAA
 CGCCGGGAGCATATCTGTTCCATGCGGCAATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTITGGCATTGTTGGACTGAAGATTTCTTCTCCAAATTCAGCTGGAAATCCAGCGGGAATCTGAGCTG
 GAGAGAAAGACGCGGACAGGCAGAAATTGAGAGAGCCGCCGGAACACGCACTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCGAAGCTCTGGGTTTCTGATCTGAAAACCTGTAAACCATAGAAAAGCTCCCGAGGAGTG
 CCTCACTCTGAGAAGAGATTACAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGAACATTTA
 CTGGGAGGTGGACGGAGGACACAAATAAAGGTGGCGGTGGAGTGTGCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGACATTGTTATTTTC
 ACATTAATATCCCGTTTATCAGCGTCTTCCCGAGGACCCACCTACAAAATAGGGGTCTTCTCGACTA
 TGAGTGTGGGACCATCTCCTCTTCAACATAAATGACCAGTCCCTATTATATACCTGACATGTGGTTTG
 AAGGCTTATTGAGGCCATACATTGAGTATCCGTCTATAATGAGCAAAATGGAACTCCCATGATCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTCTTGGCAAGGGCCCTGCAATCCAGAGACAAAGCAACAG
 TGAGTCTCTCTCAGGCAACACGCGCCTCTCTCCCGAGGGGTGAATGTAGGATGATGATCACTCCACAT
 TCTTCTTTAGGGATATTAAAGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCGCCGCAAGTGGCTTCCA
 GATGAAGGGGAGCTGGCCTGTCCAATGGAGTCAAGGTCTCATGGCTGCGCTGAGCTGGGAGGGAAGAGG
 CTGACATTACATTTAGTTTGTCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACTCTCAGGTGAAG
 AACCTCAGGAATTCCTCTCTCAGGCTGTGGTGTAGATTAAAGTAGACAAAGGAAATGTGAATATGCTTAG
 ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSCQWQVFGPKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWGPQGDLSTDSTRNDRMDHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEILDWRKKGQAEIRDARKHAVEVTLDPEAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFGAGKHYWEVDGGHNKRWRVGVCRDDVDRKEYVTLSPDHGYWVLRNGEHLTYT
LNPRFISVFPPTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPIEYPSYNEQNGT
PIVICPVTQSEKEASWQRASAIPETSNSESSSQATTFFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCTT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGCAGGAAGGCTGTGTGTCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGCCCACTAGTTTCATGGCTACTGCTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGACCCACATACCAAGAATTGACCCCTGAGCATCAGAGATGCCAGAAGTGATGCGGGGAGA
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACC GGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCAGGCACCTTGGAGTCGGCTGCCCC
AGAACTGACCTGCTCTGTGCCCTGGGCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCGGTGCCCTTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCACAC
GCCCCAGGACCATGGCACAGCCTCACCTGTGAGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTTACCCGCTCAGAAGTTGACCATGACTGCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGAGTGTGATGAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCTGACCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGCTTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGTGTGCGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCTCTCGTTGTAGTGAGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCTGGGAGATACGGGCATAGAGGATGCAACGCTGTGAGGGTTTC
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCAGACAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCCTCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTTAGAAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAGACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCTTTTATTTTTTAACTAAAAGACAGACAAATTCATA

FIGURE 86

MLLLLLLLWGRERAEQTSKLLTMQSSVTVQEGLCVHVPCSPSYPSHGWIYPGPVVHGYWFREG
 ANTQDQAPVATNNPARAVWEETRDRLFLLGDPHTKNTLSIRDARRSDAGRYFFRMEKGSIKWNY
 KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTTPMISWIGTSVSLDPSTTRS
 SVLTLLIPQPDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNSSSL
 SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNF
 LGSQQVYILNVSLQSKATSGVTQGVVCGAGATALVFLSPCVIFVVVRSCKKKSARPAAGVGDGTGLE
 DANAVRGSASQGPLETEPWAEDSPDPPPPASARSSVGEGELQYASLSFQMVKPWDSRGQEATDTE
 YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTAGCTCCAGGGCGAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
 CCAAGGAAAGTGCAAGTGAAGTACAGCAAGATTACATGAACCAACTCAGCTTCCTGCTGTTTC
 TCATAGCGACCACAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
 TCTTCGTCCTCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGGA
 TGSCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAAGACCTTCTGTGACATGACCTCTG
 GGGGTGGCGGCTGGACCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
 GGCATCGCTGGTCCAGTCAGCAGGGCAGCAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
 CACTACAACACCTTTGGATCTGCAGAGGGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
 ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
 AGAAACAGTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
 GTTTGGCATCTACCAAGAAATATCCAGTGAATATGGAGAAGGAAGTGTGGACTGACAAACGGCC
 CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAAGAAACAGCATCTTATTACTCACCTTAT
 GGCCAGCGGGGATTCACTGCGGGATTGTTTCAGTTCAGGGTATTAAATAACGAGAGAGCAGCCAA
 CGCCTGTGTGCTGGAATGAGGGTACCCGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
 GATACTTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTTCTGGTTTGGATTGGAGTGGATAT
 GGAACCTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCTCTCCCAACCATGAGATCCCAAGGATGGAGAA
 CAATCTACCCAGTAGCTAGAATGTAAATGGCAGAAGAGAAACAATAATCATATTGACTCAAGA
 AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRAWSSQQGSKADYPEGDGNWANYNTFGSRAEAT
SDDYKNPGYYDIQAKDLGIWHVFNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYSPYQGREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACCTCCAGCCTCAGAGAC
 CGCCGCCCTTGTCCTCCGAGGGGCCATGGGCCGGGTCTCAGGGCTTGTCCTCTCGCTTCCTGAGC
 CTCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGAGACAGCAATACAGGCTG
 CCTGCTCTCAGTTTACCCCCGAGGAGTATGACAAGCAGGACATTGAGCTGGTGGCCGCGCTCT
 CTGTACCCCTGGGCTCTTTGAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC
 AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCTTCTCAT
 ATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTGTCTTCTGAGTGGCCCTTCAGCTG
 TCACTGAAATGGCTTTATTGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTTGATTACCTCA
 TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG
 GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTGGAATAATTACG
 TCTTGAGTCTGGATTATCCGCATTGTATTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
 TTAAGACTTATATACAGTTTTAGGGGACAATAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLAHLVVVITLFWSRDSNIQACPLFTTPEEYDKQDIQLVAALSVTGLGFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALIV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTCTGGTCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTCCCGGCTCAGAGGACCTTGAGCGTGATGACACGAGGGCCAGCCCCGGGCTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCATGGCCAATTCCACTCTCCTAGGSGTGT
GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACACAGCCCC
CACCTTCAGGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAG
AGTTCCACAGGAACAGCAGATCTTATCGAAGCCAAGGCTCCAAAATCTTCAACTGCCGATG
GAGTGGGGAAGGTAGAACGGGGCCCGGACCTCGCTTTGCACCCACGACCCAGCAAGATCTG
CTCCCAGAGCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCTGTCTGTG
TCTACATCGCCTTCTACAGCAGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCTACTACCCATCTGGGTGACCCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCTGCCCATGCAGGAGACCATTTGGACACCGGGCAGGGAAGGGTTGGGCCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGTTCCAAAGTCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCAGAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGACGCTCAGAGGCTTTGCCACGAGGCCACAGAGAGATGCTGGG
TCCCCGAGGCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCAATGGGAGGAAGCTAAGC
CCTTGGTCTTGTCCATCTGAGGAAGATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACA
GCCTGTCACTTAGSATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAAGAGGGTGGGTGGG
GCCAGAGGAGCTCTCAGCCCTGCTAGTGGGGCCCTGAGCCCCCTTGTGCTGTCTGAGCATGS
CATGAGGCTGAAGTGGCAACCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCGAGTACTCCCCCTGTACACCCATTGCTGATG
GCACACCCATCTTAAGCTAAGACAGGACGATTGTGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACTGTGACAGCGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCAGGGTGGGGGGGGCCGGGGCCGACAGCATGTGCTGGATCTGTTT
TGTGTGCTGTCTGTGGTGGGGGAGGGGAGGGAAGTCTTGTGAACCGCTGATTGCTGACTTT
TGTGTGAAGATCGTGTCTTGGAGCAGGAATAAAGCTTGGCCCCGGGCA

FIGURE 92

MLTRCCFVFLVQGSLLVLCGQDDGPPGSEDFERDDHEGQPRPRVPRKRGHISPKSAPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHNGTGF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTFYYPG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTGCGC
 CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCGGAGCTTTCCTT
 CTGGTTGGTGTCTCTACTGATTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
 AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
 ATGTTCCGATTTCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
 AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTCTGGCTTGGGCTTTGGAATCATGA
 GTGGAGTATTTTCTTTTGAATACCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCTAT
 GGAGATTCTCCTCAATTCTTCTTTATTTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
 ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
 TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCCTTATTATGGAATAAACCTGGCG
 TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCG
 AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCCTTTTACAACCAGCGCTCCAGAT
AACTCAGGGAACCGACTTCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCGCT
 TTTTCTGAAAAATCCCTTTTTCTGGTGGAAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFPGCAFIAGFPALALVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLILFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPSTVGIRGDSQQFFLYSAFMTLVIILLHFVFWGIVFFDGCEKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCCLLCQDKNFFLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCAAAAACCTGGCTCCGGATCAGG
GAACACTACCAAAACCAACAGCAGTCAAATCAGGTCTTTCTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTGGAGCCAGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTTGTTCGGGGAGGCATCCTGCCACAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCTCCCAACTCCAGTGGCACAGATGACGACTTTCAGSTGACCACCCCT
GCAGGCATCCAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAATGGGAATTCA
GTAAGCTGTTCAAAATTTTCAACTAAGCTGCCGTAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTTCTTAAT
TACCTGAAATATTCTTGAAATTCAGAAAATATGTCTATGTAGAGAATCCCACTTTTAAAAA
CAATAATCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAARCTGGAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLLIIHSLFFGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2

FIGURE 97

GCTCAAGTGCCTTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTCTTGTCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCTGTCTTGACCATG
 GTCCCTGGCTGGCTGTGGCTGCTTTGTGTCTCCGCTCCCCAGGCTCTCTCCCAAGGCCCTAGCCTGC
 AGAGCTGTCTGTGGAGGTTCCAGAAAATATGGTGGAAATTTCCCTTTATACCTGACCAAGTGGC
 CGCTGCCCCGTGAGGGGGCTGAGGCCAGATCGTGTCTGACGGGACTCAGGCAGGCCAATGAG
 GGCCCATTTGCTATGGATCCAGATTC TGCTTCTGCTGGTGACCGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTCAACCTGGAGATGCAGGATGGACATGTCTTGTGGGTCCAC
 AGCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAAGGTGCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCGGGGTACCAGGCTTGGCATCCCTTCTCTTCTTGAAGCTTCAGACCG
 GGTAGAGCCAGGCACAGCCAACTCGGATCTTCGATTCACATCTCGAGCCAGGCTCCAGCCAGC
 CTTCGCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTTGGCCCCAGGCCCAAGGGAGC
 ACCAGGCTTGACCACGCCCTGGAGAGGACTTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCTTCAGGCCACAGGCCACTTGCACCGTGGAAATCTCCATCATAGAGAGCACTGGGTGT
 CCTTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCGCACCATGGCCCAAGTA
 CACTGGAGTGGGGTGATGTGCATATCACTTGGAGAGCCATCCCCGGGACCTTTGAAGTGAA
 TGCAGAGGGAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATCCCATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCAGCTGCTGGTG
 ATGGATGAGAATGACAACTGGCTATCTGCCCTCCCGGTGACCCCACTCAGCATCCCTGAGCT
 CAGTGCACAGGTACTGAAGTGACTAGACTGTGACAGAGGATGCAGATGCCGCCGGCTCCGCCA
 ATTCCACGTTGTGTATCAGCTCCTGAGCCTGAGCCTGAGGATGGGTAGAGGGGAGAGCCTTC
 CAGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTTCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGTGGCCATGGAACTGGCAGGCGCAGAGGGTGGCTTCAGCAGCAGCTGTGAAG
 TCGAAGTCGCACTACAGATATCAATGATCAGCCCTCGAGTTTCACTTCCAGATTTGGGCT
 ATAAGCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCTTCCGCTCATGGATTTTGCCATTGAGAGGGGAGACAGAGAAGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 CGAGCTCCAAGTCATGAGTGGTGGTGGTGGTGCAGAGTGTGGCAGAGCTGGTGGGGCAGGCC
 AGGCCCTGGAGCACCCGCCAGGTGACTGTGCTAGTGGAGAGAGTATGATGCCACCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGCGTGGCT
 CTGCATTGAGAATTTCTCGGGGAGGTGCACACGCCCGAGTCCCTGCAAGGGCGCCAGCTGGGG
 ACACCTACACGTTGCTTGTGGAGGCCCAGGATACAGCCCTGACTTTGCCCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCGCCAAGACCATGGCTTGATCTGAGTGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGCAAGCTCCTACAGCTTACCTTTGGTCCCAACCCAGCGGTGCAACGGGATTGGC
 GCTTCAGACTCTCAATGGTTCCATGCTTACCTTACCTTGGCCCTGCATTTGGGTGGAGCAGCT
 GAACACATAATCCCCGTGGTGGTTCAGCCCAATGCCAGATGTGGCAGCTCCTGGTTGAGTAT
 CGTGTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGTGGGCCGATGAAGGGCATGCCCA
 CGAAGCTGTGGCATGGGACTCCTTTGAGCGACCTGGTAGCAATAGGAATCTTCTCATCTTC
 ATTTTCAACCATGGACATGTCAAGGAGAAGGACCCGGATCAACAGCAGACAGCGTGGCCCT
 GAAGGCACTGTCTGAATGGGCCAGGCAGCTCTAGCTGGGAGCTTGGCTCTGGCTCCATCTGAG
 TCCCTTGGGAGAGGCCAGCACCCAGATCCACAGGGGACAGGACAGTGAAGAGCCCTCCCA
 TCTGCCCTGGGTGGAGGCCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACACTT
 TATCGACTGCCATGGAGGTGCTCCAAATGTACGGGTGTTTGCCCAATAATAAGCCCAAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 98

MVPAWLWLLCVSVQALPKAQPALSVVEFVPPENYGGNFPFLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVILEMQDGHVLWGPQVPLVHVKDENDQVPHFSQAI
YRRLSRGTRPGIPFLFEASDRDEFGTANSDLRFHILSQAPAQSPDMFOLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHMAQ
VHMSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIYAAPLELHVL
VMDENDNVPICPPRDPVSIPELSPFGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEFDSGHVRLRLCKNLSY
EAAPSEHVVVVVQSVAKLVGPGPGGATATVTIVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDFISRTLRFSLVNDSEGWLCEKFSGEVHTAQSLGGAQPGDITYTLVEAQDTALTALAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGYPSTILGNPTVQRDWRLOTLNGSHAYLTALAHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCGCTGGAGGAAGCCTAAGGAACCCAGGCACTCCAGTGTGCCACGCGCTG
 AGCTCAAGATTCTTCCCAGGACACAAACGTAGAGAGACCCAGCTCTCTGGAAGCACCAGCGCTTTA
 TCTCTTCACTCTCAAGTCCCTTTCTCAAGATCTCTGTCTTTGCCCCCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCACTCTGCTTTTCAGCCACAAAGAGACAGATGAAGATGCAAGAAAGAAATG
 TTCTCTCTATGTTGGTCTACTATTGCAATTAGAAAGCTGCAACAAATTCATGAGACTAGCACC
 TGTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCCAGCAGCCACCATCTCAGGGTCCAGCGTGACCTCCAAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCACCAACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCAGTGGGGCCAGCACAGTCACCA
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACAGCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATCTGAGTCCAGCACACCTCC
 AGTGGGGCCACAGACCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGTCCAGCA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCACCAATCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCATAGTCATCTACTGCACTGAGTGGGCAAGCCTGGTGGTCCCTGGT
 GCGTGGGAAATCTTCTCATCACCCTGGTCTCGGTTGTGGCGCGCTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGAGAAACAGCGCTGTCCCTGAGAAACACCTTTAAACAGCTGTCTACCACTCT
 CAGGGCTCAACCATGGCTTGGTCCAGGCCCTGGAGGGAATCTGAGGCCCCCAAGGCCCCAG
 GTGGAGTCCTAAGTGGTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCAATCTTCAGGAAGGAAGACCTGGGCA
 CCCAAGACCTGGTTTCTTTTCATTCATCCAGGAGACCCCTCCAGCTTTGTTTGGATCTCGAA
 AATCTTGAAGAAGGTATTCCTCACCTTCTTGGCTTTACAGACACTGGAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACAGCAAGAGAGAGCTGTGCTTG
 CCCCGGGTGGGTATCTAGCTCTGAGATGAACCTAGTTATAGGAGAAACCTCCATGCTGGACTC
 CATCTGGCAATCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MMQKGNVLLMFGLLLHLEAATNSNETSTTSANTGSSVSSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFTASSGISIATNSESSTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASATNSESSTLSSGASTATNSDSSTSSGASTA
TNSESSTSSGASTATNSESSTVSSRASATNSESSTSSGASTATNSESRTSSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTSSGASTATNSESSTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTSSGANTATNSESSTVSSGASTA
TNSESSTSSGVSTATNSESSTSSGASTATNSDSSTSSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEI FLITLVSVVA
AVGLFAGLFFCVRNLSLSLRNTFTAVYHPHGLNHGLGPGGNGHAGPHRPWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCGGACGCCTCGCGTTACGGGATGAATTACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGGCCACGCACCCCACTCGGCGTCGCGCGCGTGCCTGCTTGTACAGGTG
GGAGGCTGGAATATCAGGCTGAAAAACAGAGTGGGTACTCTTCTGCGGAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAATTTGGTGTCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTATGGAAGGGATTACCTTATACTGACTCTGTTTTGGGGAAGCTTTTTGGAAGCATTT
TCATGCTGAGTCCCTTTTACCTTTGATGTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTAITGGAGACCATGTTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTTGTTCCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCCTCAAAGCAGTCTCAAAGGTGTTCTCGGATTTGGTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTGGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAGGACTGATCTCACAGAAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGTAGACCGTCTAAGAGAAGGTAAGAACCCTTGATGCTGCCATG
ATATCACTGTGGGTATCCTCACAACTTCTCAATCAGAGAAGCACCTCCTCCAAAGGAGACTTT
CCGAGGGAATCCACTTTCAGCTCCACCGGTATCCAATAGACACCTCCCCACATCCAAAGGAGGA
CCTTCAACTCTGGTGCCACAACCGTGGGAAGAGAAGAAGAGAGGCTCGCTTCTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTTGCAAGCTGAACTCAGGGTC
CTTGTGGTCAAATGCTCTCTATACTGTATTGGACCTGTTGAGCCCTGCAATGTGCTACTCAT
ATATTTGACAGTCTTGTAAAGTGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCAGACTTTTACACAAACAGGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGTTTGGCATGTGAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTCTAAGCTCAGATGCATTTTGTGATGACTATGTCGAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGACTTAAATTTTGGGAAAAATATTGCTACAATTTT
TTTAATCTCTGAATGTAATTCGATACGTGTACATAGCAGGAGGTGATCGGGGTGAAATAACTT
GGGCCAGATATTAATAACAATCATCAGGCTTTTAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTFWGSFFGSIIMLSPL
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDMM
FLWNCIMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFFEGTDLTENSKRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFREIHFFVHRYPIDTLPTSKEDLQWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEACYRLHLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCTGCGCCTGAGACAGCTGGCCTGACC
TCCAATCATCCATCCACCCCTGCTGCTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCAATGGCTTTTGTGCTCATTTTGGTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGAGCCGGGCAAGTTTGTCCAGGCTTGGTGGGGAGGAGCCCGCTTCTCCTGCTCCCTCT
TTCCTGAGACCAAGTGCAGAGGCTATGGAAAGTGGGTTTCTCAGGAATCAGTTCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAATCGA
GTTTGTGAGGACTCCATTGACGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCTGTATGGGTGCTGGTTCCAGTTCACAGATTACAGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCTGCGGATATGTTAGCGGAGGTATCCA
GTTACTCTGCTCTCTCAGGCTGGTTCGCCAGCCACAGCCCAAGTGGAAAGGTCACAAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGGAGCATATTGTTTCCATCCACTTGTGTAGCAGAGATCATGA
GGTGGAAATCCAAGTATTGATAGGAGAGACGTTTTTCAGCCCTCACCTTGGCCCTGGCTCTA
TTTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGTGTGTGAGGGATGATAATTGTTTTCTC
AAATCCAAAGGGAAAATCCAGGCGGAACGGACTGGAGAGAAAGCACGGACAGGCAGAAATTGAG
AGACGCCCGGAAACACGCAAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGG
TTTCTGATCTGAAAATCTGAACCATAGAAAAGCTCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
TTTACAAAGGAAGAGTGTGGTGCTTCTCAGGGTTTCCAAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTGGGAGTGTGTGGGATGACGTAGACAGGGGGAGAACA
ATGTGACTTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAAACAGAACTTTGATTTTC
ACATTCAATCCCATTTTATCAGCCTCCCCCAGCAGCCCTCTTACACAGTAGGGGTCTTCTCT
GGACTATGAGGGTGGGACCATCTCTTCTTCAATACAAATGACCAGTCCCTTTATTTATACCTGTC
TGACATGTCAAGTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACAGGGAAG
GGGACTCCCATATTCATATGTCCAGTGTCTGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCACACACAGACCCAGACACAGCCAAAGGAGAGTGTCCCAGCAGGTGGCCCACTCTCTCT
CCGAGCCTGCGCACAGAGAGTCAAGCCCCCACTCTCTTTAGGGAGCTGAGGTCTCTCTGCC
TGAGCCTTGACAGCGGCGAGTCAAGCTTCCAGATGAGGGGGGATTTGGCTGACCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGACGGAATAGATCACATTAGGTTTAGTTTGTGAAAA
CTCATCCAGCTAAGCGATCTTGAACAAGTCAAACTCCAGGCTCCTCATTGTCTAGTACGG
ACAGTGATTCCTGCCCTCAGAGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTGCAAGTT
TGAGGGCAGTGTGTCTAATGATGTGTTTTATATATACATTTTCCACCATAAATCTGTT
TGTCTATTCCACATTAAATTTACTTTTCTTATACAAATCACCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAAATAAGAGGAGGTAGGATTTTCTACTGATCTATAAGCCAGCAT
TACCTGATACCAAAACACGCAAGAAACAGAAAGAGAGAGGAAAACTACAGGTCCATATCC
CTCATTAACACAGACAAAAATTTCTAAATAAAATTTTAAACAAATTAACATAACATATATTTA
AAGATGATATATAACTACTCAGTGTGGTTTGTCCCAAAATGCAGAGTTGGTTAATATTTAAAT
ATCAACAGTGTAAATTGACGACATTAATAAGTAAAAAGAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGNFPQPTAKWGPQGQDLSSDGRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLOGALCGVVMGMIIIVFFK
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPEAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDGRGNVTLSPNNGYWVLRLLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGAGCTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAGAG
TTTGGTTGGAAACCCCTGGGTTATCGGCCCTGGTCATCTTCATATCCCTGATTGCTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCAATTACAAAC
TGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCAACAATTTTACAGAAATGAGCCAGAGACTTGAAAT
CAATGGTGAAAAATGCATTTTATAACTCCATTAAAGGGAAGAAATTTGTCAGTCTCAGGTTATCAAGTTC
AGTCAACAGAGCAATGGAGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGACTCCTGA
AACTGTAGNTAAATTTGTCAACTGTGTTTACATGAAAGCTGCCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGACAGAAGTAGANGAGGGTGANTG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAAATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATPAAAGAACCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAAACCTTCGAAATGAAACGGGGTCTCGGAGAAATAAATTGCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTGCGAGAGCTTTCTAGCCCTGTTCCTACACAAATGCAGTACATAGAGTTTGTC
TCCCTGATGCATCCTATGAGTTTCAACCCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTTCGACPAGCACAGTGACTCTCATAGACGCTACAACTTGCAATGAACC
TCAAGCTTACAAATGACGCCATACTCCTAGAATGTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGCCACTGGTTAGTTTCAGATGCTAGAGATATCTGTAACCTTGCTGGAATAGTG
AGCTGGGGAGTGAATGTCCGAAACCCAAAGGCTGGTGTATTACTAGAGTTACGGCCTTGCGGGACGTG
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAGCCTCATGGACAGATAAACATTTTTTTTGTGTTTTG
GGTGTGGAGGCCATTTTAGAGATACAGAAATGGAGAACTTGCAAACAGCTAGATTGACTGATCTCA
ATAAAGCTTTGCTGATGCATGTATTTCTTCCAGCTCTGTCCGACGTAAGCATCTGCTTCTGCCA
GATCAACTCTGTCACTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTGTTCTCTAGAAATTTTGTGAGAAATTTGACTTGTGACATAAATTTGTAAAT
GCATATATACAATTTGAAAGCACTCCTTTCTTCAGTTCTCAGCTCCTCTCAATTCAGCAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAGAAAAATATAAGAAAGAAAAAATCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGAATATTAGAATGATCATATTCAATTATGAAGGTCAAGCAANGACA
GCAGAAATCAACTCACTTCATCTTTAGGAAGTATGGGAAGTAAAGTAAAGGAGTCCAGAAAGAGGCCAAG
ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT
ATAAATATATACAACCTCATGCAATGTACTTGTCTANGCAAAATTAAGCAAAATATTATTTAACTATG
TTACTGAGGATGTCACATATACAAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYSTLSFTTDKLY
AEFGREASNNFTMSQRLSEMVKNIFYKSPLEEFVKSQVIFKSQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVFYTNVHRVCLPDASYEFQPGDMFVTGFGALKNDGYS
QNHRLQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAGCGCTCTCCAGCTGAAGCCAAATGCAGCCCTCCGGCTCTCCGGAAGAAGTTCCCTGT
 CCCCGATGAGCCCCGCCCTGCGTCCCCGACTATCCCCAGCGGGCGTGGGCGACCGGGCCCGCAGC
 GCCGCGATCGCTGCCCTTTTGGCCCTTGGGATAGGATGTGGTGAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACATGGCCAGAGAAATTCGCTGAAGTGTCTGCGCTGCCCTGCTTACGCGCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGAACACCTAA
 ATAAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAAGGAGCAGTCATTTGACTTACTTTCTC
 GTGGTTCACTCCGCTCAGATTGCTGTTTGGCTTTTCCCTTATCATTTGGGGATGTTAGGATATTG
 TGGAAACGTTGAAAGAAATCTGTGCTTCTTGCATGCTACTTTGGAACTTTCCTCTCATTTTCT
 GTGTAGAATGGCTTGTGGCGTTTGGACATATGAACAGGAATTTATGGTTACCTAGATATCGGTGGCTTACTCA
 GATATGGTCACTTTGAAGCCAGGATGACAAATTAATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTTTTTCAGAGAGAGTTTAAATGCTGTGGAGTAGTATATTTCACTGACTGGTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTTAGAGAATTTCCAGGAATGTTCCAAA
 CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTTGGGAAGAAAATGTTATCCCTT
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTTGCCATGATTCTCACATTACTTGTCTCTGGGCTCTGATATGATAGAAAGGAGCGTGGG
 ACAGACCAATGATGTCTTGAAGAATGACAACCTCAGCACTCTCATGTCCTCAGTAGAACT
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATAAAAGAAAATGTCACAGAAGAAAACCAAACTTGTTTATTGGACT
 TGTGAATTTTGTAGTACATCTATGTGTTTACAGAAATATGTAGAAATAAAAATGTTGCATAAAA
 TAACACCTTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTCATGTCAATAAGC
 ACCACCTGGACAATAATTTGATGCCCTTAAAATGCTGAAGACAGATGTCAATACCACCTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGTATAGCATTTCCGCA
 TCCATGCAACAGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
 CTAGTATATAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTCTGTATCTAAATAAATATATATACAGAAACTTTCATATTTGGTGACT
 ACCTAAATGTGATTTTGGCTGGTTACTAAATATCTTACCACCTTAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGATTTTTTGTATATATAGTCTGTGTTAAATCTGTATATAATTCAGTCGAT
 TTCAGTCTGATAATGTTAAGAATAACCATTAAGAAAGGAAAATTTGCTGTATAGCATCAAT
 ATTTTATGCCCTTCTGTATAAAGCTTTTACTATTCTGTGCTCGGGCTTATATTACACATATAAC
 TGTATTATAATACTTAACCACCTAATTTTGAATAATTACCAGTGTGATACATAGGAATCATTTATC
 AGAATGTAGTCTGTCTTTAGGAGTATTATAAGAAAATTTGCACATAACTTAGTTGATTGAGAA
 AAGGACTTGTATGCTGTTTTTCTCCCAATGAAGACTCTTTTGGACATAAACACTTTTAAAAAA
 GCTTATCTTTGCCCTTCTCCAAACAGAAAGCAATAGTCTCAAGTCAATATAAATTTCTACAGAAA
 TAGTGTCTTTTCTCCAGAAAATGCTTGTGAGATCAATAAAACTGTGACAAITTTAGAGATT
 CTTTGTTTTATTTCACTGATTATATACGTGGCAATTTACACAGATTATTAATTTTTTACAA
 GAGTATAGTATATTTATTGAAATGGGAAAATGCAITTTTACTGTATTTTGTGATTGTTTAT
 TCTCTAGAAATGGAAGAAAATAAAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAMMRDYLNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWVTEQELMVPVQNSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSLDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAAGCCAGAGCTGTGGAACCTTTATCCCACTATCCTCATCCTCTTCTCTGATAAAGCCCTACCAGTGC
GATAAAGCTTTTCTCTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAGAAGGGGACAAAGGAACA
CCAGTATTAAAGGATTTTCAGTGTTCCTGGCAGTTGGTCCAGAAGGATGCTCCTCATTCTCTCTCACTG
CCTCTTCATCAGGACCTCCGTGTCACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCAACAGTTGGATGAGTCTCAAGTCCCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACCTTACGGGCATGGCGGAGATGCCATGCTACTTCTGCATACCAAGAAACCACTGTGGAAACCA
CGCACCTGCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCAATTGTGCACGCCAGGCTTGTGCCAGCT
TCAATGGGAACCTGCTGTCTGGAAACACACGGTGAAGTCAAGGCTCCCTCGAGGGCTACTATGTGTATCGT
CTGACCAAGCCACGGCTGTCTTCCACGTCTACTGTGTCAATTTTATGACATCTGGCAGGAGCTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGGCTCCAGGAAGTGTCTAGGCCCTGACAGCGAGACATGCTTTG
ATGAAATGAATGTGACCAAAACACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCCGTGTGCTAAGAAATGATGGCAAGACTTGTGAAGACTTGAAGATGCCAATACAA
TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAGGCTACCGTGTGAATGTCCCGGGGCTGTGGTCTGT
CTGAGGATTAACCACTTGCCAAGTCCCTGTGTGTGCAAAATCAAAATGCCATTGAAGTGAAATCCCGAGGGAG
CTGTTGTGTGGCTGGAGCTCTTCTGACCAACACCTCTGCGGAGGAGTGTCCAAAGGACCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAMTGAAGAATTTGGCCAGCAACCTCGTGA
CAGTCTACCCAAAGCAGACCCGGGGAGCAGCGGGGACTTCATCATCCGAAACAGCAAGCTGCTGATCCCGGTG
AAGTGCAGTTTCCACGCCGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCACTGGAAAT
CATGAGGCCAAATCATGGGATCTTCCATTCACTCTGAGATCTTCAAGGACAATGAGTTTGAAGAGCTTACC
GGGAAGCTCTGCCACCCCTCAAGCTTCTGTGACTCCCTCACTTTGGCATTGAGCCCGTGGTGCAGCTGAGCGGC
TTGGAAGAGCTTGGTGGAGAGCTCTTTCGCCACCCCACTTCCAAAGATCGCAGAGGTCTGAAATCACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGATCACCTAGCAAGCACTTCCAGG
TCCCTGTCTTCAAGTTGTGGGCAAGACCACAAGGAAGTTCTTCTGCATGCGGGTCTTGTCTGTGAGATG
TTGGAAGAGCTTCCGCTGTGCCAGGCTTGGCAGCGGCAATCGCTGTGGGGCAGGAGGAGAGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGGGCCGATCGGCATCGACTGGGAGGACTAGTCTGTAGCCATACCTC
GAGTCCCTCAGTTGGACGCTCTGCTCTTTGGAGCTTCTCCCGCCACCGCCCTCAAGAACATCTGCCAACACG
TGGTTCAAGCTTCACTGTGAGTTCAGACTCCCAAGCACAACCTCACTCTGATCTGTGTCATTCAAGTGGGCA
CAGGTCACAGCACTGCTGAACATGTGGCTTGGTGGGTTTCACTTTCTAGGGTTGAAACTAACTCTGTCA
CCAGAGAAGCACTCACCCATTTCCTCATTCTCTTACACTTAAATACCTCGTATGGTGGCAATCAGAC
CACAAATCAGAAGCTGGGTATAATATTTCAAGTTACAACCCTAGAAAATTAACAGTTACTGAAATTAATGA
CTTAAATACCAATGACTCCTTAATATGTAAATTAAGTTATACCTTGAATTTCAATTCAATGCAAGCTAA
TTATAGGGAATTTGGAAGTGTATCAATAAAGCAGTATATAATTT

FIGURE 110

MPFFLLLTCLFITGTSVSPVALDPCSAIYISLINEPWRNTDHLQDESQGPPLCDNHVNGENYHFTGMAGDAMP
TFCLPENHCGTHAPVWLNGSHPLEGDGIVQRACASFNGNOCCLWNTTVEVKACPGGYVYALTKPSCFHV
YCGHFYDICEDECHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYACECGVGRV
LASDGKTCEDVEGCHNNNGCASHCLGSEKGYQCECPRLVLSEDNHTQVPVLCKSNAIEVNIIPRELVGG
LELEFLINTSCRGVSNQTHVNLFLSLKTCGTVVDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYPVNLNSPFLIMSRNHGIFPFTLEIFKDNFEFEPYREALPTLKLADSLYFGIEFVVHV
SGLESLVESCFAFTTSKIDVLKYLLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDKHEVFLHCRV
LVCGVLDERSRCAQGCHRRMRGAGGEDSAGLQQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZF domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGCAAGGATGCTGGGCGTAGGGACCAAGGCGTGCCTGCACTCGG
 GCCTCCTCCAGCCAGTGTGTGACCCAGGACTTCTGACCTGCTGCCAGCCAGGACCTGTGTGGGAGGCCCT
 CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCAGTGT
 TACAGGATCCTGCAGTGATCAACCTCTGAACAGGCTCGATGTCAAACCCCTGGCGAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCTGGCGAGTATCATCATTTGT
 GGTGTCTCCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCGAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAACCTGTTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGTGTAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTGTGTAATCACAGAAAAACGCCAG
 GAGCTTCGCATGCCGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCTTGCACCTGTCTTGCCTG
 TGGGAAGAGCCTGAAGACCCCCGTGTGTGGGTGGGGAGGAGGCCCTGTGTGATCTTGGCCCTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCACTCTGACCCCACTGGGTCTCAGCGCA
 GCCACTGCTTCAGGAACATACCGATGTGTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
 CTTCCCATCCTGGCTGTGGCCAAAGATCATCATATTGAATTAACCCCATGTACCCCAAGACAAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCTCCACTCACTTCTCAGGCAGAGTCAGGCCCATGTGTCTGCCCTCTTT
 GATGGAGAGCTCACTCCAGCACCCCACTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAA
 GATGCTGCACATACTGCTGCAGGCTCAGTCCAGGTCTTGTGACACACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAGATGATGTGTGCAGGCATCCCGAGGGGGGTGTGGACACTGCCAGGCT
 GACAGTGGTGGCCCTCATGTGTACCAATCTGACCACTGGCATGTGGTGGGCATCTGTAGCTGGGGCTATGG
 CTCGGGGGGCCGAGCACCCAGGAGTATACCCAAAGTCTCAGCCTATCTCACTGGATCTACAATGTCT
 GGAAGCTGAGCTGTATGCTGCTGCCCTTTGCACTGCTGGGAGCCGCTTCTTCTGCCCTGCCCACT
 GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAGGGCCCTCAATTCCTGTAGAGACCCCTCGCAGCCAGAGGCCCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCACACTTGGTGTCTCCAGCATCCAGGGAGAGACACAGCCCACTGTACACAGGTCT
 CAGGGGTATTGCTAAGCCAAGAGAACTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAGCC
 CAGATCACTGTGGCTGGAGAGGAGAGGAAGGGCTGTGCGCCAGCCCTGTCCGTCTTCACCCATCCCAA
 GCCTACTAGAGCAAGAAACAGTTGTAAATAAATGCACCTGCCCTACTGTTGGTATGACTACCGTTTACT
 ACTGTTGTCTATTGTATTACAGCTATGCCACTATTATTAAGAGCTGTGTACACTCTCTGGCAAAAAAAA
 AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYFLOG
 QFLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
 FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
 KTPRVVGGEEASVDSWFWQVSIQYDKQHVCGGSILDPHWLTAHCFRKHTDVFNWCVRAGSKL
 GSFPSSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
 FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMCAGIPEGGVDTCCGDSGGPLMYQS
 DQWHVVGIVS#GYGCGGPSTPGVYTKVSAYLNNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCTCAGCCTCCCAAGGTGCTGTGATTA
TAGTGTAAGCCACCGTGTCTGGCCCTCTGAACAACCTTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACCTGTAGGATTCTGACTATGCTGTGGCTAGTGTCTCTACTCTACCTACATTAAAAATC
TGTTTTTTGTTCTCTTGAACCTAGCCTTTACCTTCTTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAACCTGACCTTCACTCTGGAACGAGAACAGAGTTTCTACCCACACCGTCCCTCGAAG
CCGGGACAGCCTCACCTTGTGGCTCTCGCTGGAGCAGTGCCTCACCACTGTCTCACCTCT
GGAGGCATGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGTGTGGCTTTCAAGGTGGG
CTTGCCCTGGCGGTAGAAGGGATTGCAAGCCCGAAGATTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACTGTCTCA
GAAGCAGTGGGTGAGACATCAGCTGCCCGCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACCTGTGCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATCTCCTGGAACATGAGG
GAACGCCGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGCCAAATATGGGTGAGAAAAGATG
GAGTGTTGGGTTATCACAAGGCATCAGTCTCTGCATTGAGTGACATGTGGGGGAAGGGCTG
CCGATGGGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTCCGCCCGGAT
CCACGTACCGCTGCTGAAGGGCACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGCTGCGGAAGGAGCAAGCAAAGTGACCATTCTCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAGGCCACCAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTGAGCAAGGCCGTGAGAGCTGATCAGAAGGCCCTGT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGTGCAAAATCCCAGGCCAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTTGAGCTGTCCCAAGACCAGAGGAGCTAGAGCTTGTT
CAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAACCCCAAAATCCA
AACCTAAGAACAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCTGTAATGCCAAT
TTTGGGAGGCCGAGCGGGTAGATCACCTGAGTTCAGAGTTCAAGACCAGCCTGGCCAACTGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGSTATTGTAA

FIGURE 114

MLWWLVLLLPTLKSVCFLVTSLYLPNTEDLSLWLWPKPDLSHGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVOVAEPLGSCGFOGGPCGRRRD

Signal peptide:

amino acids 1-15

[illegible]

FIGURE 115

CAGCAGTGGTCTCTCAGTCCCTCCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
 AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAAGAAA
 ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCCT
 GTTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACTT
 TCTACAGCAATGGAGAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
 TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
 CATCTACTTCGTGGTCTTCAAAAATGTTTTATCAAAACTCAGATTAAGTGATTCTCTGAATTTT
 CTGAACCAGAAGGAGAAATAGATGAGAATGAAGAAATTACCACAACCTTCTTTGAACAGTCAGTG
 ATTTGGGTCCAGCAGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAATTCAAAATTCCTGGA
 GATTGTGTATACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTTCTGAGTTACAAG
 ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCTGCCCACGAAAAAAGGGATTGAACAAAT
 GAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCCTCACGCCAGACAAGCAAGTGAGGA
 AGAACTTCCAATAATGACTATACTGAAAATGGAATAGAAATTGATCCCATGCTGGATGAGAGAG
 GTTATTGTGTATTTACTGCCGTCGAGGCARCCGCTATTGCCCGCGCGTCTGTGAACCTTTACTA
 GGCTACTACCCATATCCATACTGTGTACCAAGGAGGACGAGTCATCTGCTGTCTCATGCTCCCTG
 TAACGTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAAATAGAGGTTTGAGCTCAAATGCTTAAAC
 TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCATGAGGCACTGGCCCTT
 GGTAGGCAGCTCTCCAGAATTACTGTAGGTAATTCCTCTCTTCATGTTCTAATAAAGCTTCTACA
 TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVPGILALTLLIVLFWGSKHFWPEVPPKAYDME
HTFYNGEKKKIYMEIDPVTREIFRSGNGTDETLEVHDFKNGYTGIFYVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPAEKPIENRDFLKNKILEICDNVIMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGRNYCRRVCEPLLGYYPYPICYQGGRVICRVIMPCNWWVARMGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCOCTCAGGAGCGOSTTACGTTTACACACCTTCGGCAGCAGAGGGCGGCAGCTTCTCGCAGGGGCA
GGGGGGGGGGCAGGATCAGTTCACCAACACATGCCAAGTGGTGGCGTCTCTCTGTCACCTCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCCTCAGCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGGGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTTACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTGGAGGCCCTGATGATCTGTAGGCATCGTCT
GGGTGCCATTGGCTCTCTGATCCATCTTGGCTGAAATGCATCGCATTTGGCAGCATGGAGGACTCTG
CCAAAGCCAACTGACACTGACCTCGGGGATCMTGTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCACATGCTGTTGACTAACTTCTGGATGTCCACAGCTAACTGTACACCGGCATGGGTGG
GATGTTGCAGACTGTTCAGACCAGGTACACATTTGGTGGGCTCTGTTGTGGGCTGGTCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCTGGCACCAGAAAGAACCAATACAAA
GCCGTTTCTTATCATGCTCAGGCCACAGTGTGCCCTACAAGCTTGGAGGCTTCAAGGCCAGCATGGCTT
TGGTCCAAACCAAAAACAAGAAGATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATC
CTTCCAGCAGCAGCTATGTGTAAATGCTCTAAGACCTCTCAGCACGGCGGAGAAACTCCCGAGAGCTCA
CCCAAAAAACAAGAGATCCCATCTAGATTCTTCTGCTTTTGTACTCAGCTGGAAGTTAGAAAGCCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGOCTCAGTCTCTGCTCTAAATATTCACCAATAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCAGATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGATGTGGTTTAACTCTCTCTCAGATTTTATGATGATTAGACAGATCCCCCTC
TTCTCTCTAGTCAATAAACCATTGATGATCTATTTCCAGCTTATCCCCAAGAAACTTTTGAAGGAAA
GAGTAGACCCAAAGATGTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATTTTGAATCTCTCCAGCCCATGATCTCGGTTTCTCT
ACACTGTGATCTTAAAGTTTACCAAAACAAAGTCATTTTCAGTTTGGAGCAACCAACCTTTCTACTGCTG
TTGACATCTCTTATTACAGCAACCACTCTAGGAGTTCTCTGAGCTCTCCACTGGAGTCTCTTCTCT
GGCGGGTCAGAAATTGCTCCCTAGATGAATGAGAAAAATTTTCTTAAATTAAGTCTTAATATAGTTAA
AATAAATAATGTTTGTAGTAAATGATACACTCTCTGTGAATAGGCTCACCCTTACATGTGATAGAG
GAATGAAAAAATAATGCTTTGACATTTGTCTATATGGTACTTTGTAAAGTCATGCTTAATGACAAATTC
ATGAAAGCTCACAOCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAAGCCAGAGT
TCGAGACTAGCTGGGCAACATGGAGAGGCCCTGTCTCTCAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCAGCATTCOSGAGGCTGAGGTGGGAGGATCACTTGAAGCCAGGAGGT
TGGGGCTGAGTGAGCCATGATCACCACTGCATCCAGCCAGGTGACATAGCGAGATCTGTCTAAAAA
AATAAAAAATAATATGGAACACAGCAAGTCTTAGGAAGTAGGTTAAACTAATCTTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMNSTQDLYDNFVTSVFQYEGLRSCVRQSSSGFTECRP
YFTILGLPAMLOAVRALMIVGIVLGAIGLLVSI FALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYIHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTCTGTGGCAGAGAACCTGCTCAAAGCAGAAGTAGCAGTTCGGAGTCC
AGCTGGCTAAACTCATCCAGAGGATAATGGCAACCCATGCCTTAGAAATGCTGGGCTGTTTCTTG
GTGGTGTGGAAATGGTGGGCACAGTGGCTGTCACTGTATGCCTCAGTGGAGAGTGTGGGCTTCATT
GAAAAACACATCGTGGTTTTTGAAACTTCTGGGAAGACGTGGGATGAATTCGCTGAGGCAGGCTAA
CATCAGAGTGCAGTGCAAAATCTATGATTCCTGCTGGCTCTTCTCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCGCTGATGTCCTTCTGGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTCTTATA
ACTCAATAGTGAATGTGCCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACACGGCA
CTGCTGCTGATTTGTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGTGCAACGAAAAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAGTTATCACACCGAAGAAGTCAACGAGCG
TCTACTCCAGAAGTCAGTATGTGAGTTGTGTATGTTTTTTTAACTTTACTATAAAGCCATGCAAAAT
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAAAGTTGATTTACTGTCTTAACTGCCT
AATCTTAATTACAGAACTGTGCATCAGCTATTTATGATCTATAAGCTATTTACGCAGAAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAATATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTTTACTTCAAAATGACATGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGAA
ATCATGGATAGGTTGAAGAAGTTACTATTAATTTGTTAAAAACAGCTTAGGGATTAAATGCTCCCA
TTTATAATGAAGATTAAAAATGAAGCTTTAATCAGCATTGTAAGGAAATGAATGGCTTCTCGATAT
GCTGTTTTTACGCTAGGAGTTAGAATCCTAACTCTTTATCTCTTCTCCAGAGGCTTTTTTTTTT
CTTGCTGATTAATTAACATTTTAAACGCAGATATTTTGTCAAGGGCTTTGCATTCAAATGCTT
TTCAGGGCTATACTCAGAAGAAAGTAAAAAGTGTATCTAAGAAAAGTGATGGTTTTAGGAAAGTG
AAAAATTTTTTGTGTTTGTATTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGATGAGT
ATATTTTAAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGACGACGAAAAATA
TGCTCTGGTTTTTCATTTGCTTACCAAAAAACACAACAAAAAAGTTGCTCTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTTGTCATTTTGTCTGTGAAAAATAAATTTCTCTCTGTGA
CCATTTCTGTTTGTGTTTACTAAAACTGTAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAA
ACTGACAATCCAATTTGAAGTTTGTGTCGAGCTCTGCTAGCTTAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATGTCATATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMFPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLALLSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDfYNSIVNVAQKRELGEALYLGWTTALVLIIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCCGCGGGGTGAAAGCCGATTGATGCAGCTCGCGGGCTCGGAGCGCGCGGAG
CCAGACGTGACCACTTCTCTCTCGGTCTCTCGCCTCCAGCTCCGCGCTGCCCGCGAGCC
GGGAGCCATCGGACCCAGGGCCCCGCGCCTCCCCGACGGCTCCGGGCGCTCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCTCGAGCGCTCTGAGATCCCCAAGGGGAAGCAAAAGCGCAG
CTCCGGCAGAGGGAGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGCTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCCGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAGCTTTGAGGAGTCTGGACACCOCACTAC
AAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAATTCGGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCACTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTACAGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATCAACAATTAATATTATCG
CACTTCTCTGTGGAAGGACTTTTGAAGGAATTGGTCTGGATTAGTGGAATGTTGCTATCTGGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAATAACAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACITAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTTATCATTTTG
CTTCAATCAAAAGTGGTTTCAATATTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGGTCTTTTGTGTTTTCTCTTAGTATAGCAATTTTA
AAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTATATCTGT
TAAATAAAANTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKKGKQKQLRQREVVDLYNGMCLQGPAQVPGR
DGSFGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGSLPIEAI IYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSR IIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCTGCCTCTGGGCTCCAAAGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAAGCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATGTAGGATCCAGTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAACCTGTAAGATGCAAAAACGTAATATCCATGANGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTCTCGGTGGTGGCAAGAAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGCTCAATTTT
 TCTTCTCGGTGCTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCTAAGCCAAAGCAAAAGACCTAAGGACACCTTTGAACAATACAAAGGATGCGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAATTTACAGAGGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTGGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACTGGCTATAC
 CTGACCATAACCATATCAGCAATATTGACGAAATGCTTTAATGGAATACGAGACTCAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAAACAATACCTTCAGACTGTGACAAATT
 TACGAACTTGGATCTGCTTATAATCAGTGCATTTCTGCGGATCTGAACAGTTTGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAGAACCTTCACTGGAGCACATCAATTTTCCAAAGCTCAAC
 CTGGCCCTTTTTCCAAGTTTGGTCAAGCTTCAAGACCTTTACTTGAAGTGAATATAATCAGTGT
 CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTTACAAGGCTTGATTATCAGGCAATGAGA
 TCGAAGCTTTAGTGGACCCAGTGTTTTCCAGTGTGTCGGAATCTGCAGCGCTCAACCTGGAT
 TCCAACAAGCTCAATTTATTGGTCAAGAGATTTTGATTCTTGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGTCTCCCTTGTAACTGGCTGAAAGTT
 TTAAGGCTTAAGGGAGAATAACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAATGTG
 ATCGATGCAATGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGAATCGGCCAG
 GGTCTTCCCAAGGCCGACGTTTAAAGCCCAAGCTCCCAAGGCTGAAGCATGAGAGCAAAACCCCTT
 TGCCCCCAGCGTGGGAGCCAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGCGAGCTGGCGCTTTTCTGTGCGTGTGCTCATCTGCTGTTTAT
 CTACGTGTATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAGAAAAGACAGTCCCTAAAGCAATGACTCCGACACCCAGGAATTTATGTA
 GATTATAAACCCACCAACCGGAGACCAGCGAGATGCTGTGAATGGGACGGGACCTTGACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAGAGCTCTTAAAGCT
 GGGAAATAGTGGTGTCTTATTGAACCTGGTACTATCAAGGGAACGCGATGCCCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCTTCTTTGTCCGTTTATGTGCAATCATATACT
 GGTCAITTTCTCTCATACATAATCAACCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCGGTTTAAATAAATACCTATTGTATAAGACCTTTACTGATTCCATTATGTGCAATTT
 GTTTTAAGATAAACTCTTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCGKMYCESQKLQEIIPSSISAGCLG
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLNLNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNLAFFPLVSLQNLVYLQWNKISVIGQTMSTWSSLRQLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLTFIGQEILDWSISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRHRHKKKQSLKQMPSTQEYVDYKPTNTETSEMLLNGTGPCCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATCGTCCGTCCTCCGAGGAGGAGAGGCTTTTGCCGCTG
 ACCCAGAGATGGCCCCGAGCGAGCAAATCTCTACTGTCGCGCTGCGCGGTACCGTGGCCGAGCT
 AGCAACCTTTCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
 GGTGGGAGACGCTGCAAGAGAATCTGCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
 ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACAGTAGT
 GTATTCTGGAGTCCGAATGGTCACATATGAACATCTCCGAGAGGTGTGTTTGGCAAAAGTGAAG
 ATGAGCATTATCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
 TTAGCCAATCCAACCTAGTACCTAGTGAAGTTTCAGATGCAATGGAAGGAAAAAGGAACTGGAAG
 AAAACCATTCGATTTCTGTTGTACATCATGCATTTGCAAAATCTTAGCTGAAGGAGGAATAC
 GAGGGCTTTGGCAGGCTGGGTACCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
 ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
 TCACGGTTTATCAGTTTATGTTCTGAGCTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
 TCAAAAGCAGATAATGAATCAACCACGAGATAAACAGGAAGGGGACTTTTGTATAAATCATCG
 ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTTCATGAGTCTATATAAAGGCTTTTACC
 ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
 TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVFEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
 APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
 VIGGMMAGVIGQFLANFTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
 NIQRAALVNMGLTTYDTVKHYLVLTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQF
 RDKQGRGLLYKSSTDCLIQAVQGEFGMSLYKGFLPSWLRMTFWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGCGCGGCGCGGACGAGCGCGCGGTGAGCTCTCGAC
 CCCCGTGTGCGGCTAGTCCAGCGAGCGGAGCGGCGCGGTGGGCCCATGGCCAGGCGCGGCATGG
 AGCGGTGGCGCGACCGGCTGGCGGTGGTACGGGGGCTCGGGGGCATCGGCGCGCGGTGGCC
 CGGGCCCTGGTCCAGCAGGGAAGTGAAGTGGTGGGCTGCGCCCGCACGTGGGCAACATCAGAGGA
 GCTGGTCTGAATGTAAGAGTGCAGGTACCCCGGACTTGATCCCTACAGATGTGACCTAT
 CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTAGACATC
 TGCATCAACATGCTGGCTTGGCCCGGCTGACACCTGCTCTCAGGCAGCACAGTGGTTGGAA
 GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATGTCACACGGGAAGCCTACAGTCCATGA
 AGGAGCGGAATGTGACGATGGGACATCATTACATCAATAGCATGTCTGGCCACCGAGTGTTA
 CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGGGGACT
 GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCAGGTGTGGTGG
 AGACACAATTGCGCTTCAAACCTCCAGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
 ATGAAGTGTCTCAAACCCGAGGATGTGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCACA
 CATCCAGATTGGAGACATCCAGATGAGGCCACGAGGACAGGTGACCAGTACTGTGGGAGCTCC
 TCCCTCCCTCCCAACCCTTCATGGCTTGCCCTCCTGCCCTCTGGAATTTAGTGTGATTTCTGAT
 CACGGGATACCACTTCTCTCCACACCCGACCGAGGGGTAGAAAATTTGTTTGAATTTTATA
 TCATCTGTCTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAGGAGTGGTGTCCC
 TAAATGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCACCTTGGCCTTGTCTGCTCTCAGT
 TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCATCTTCTTGACCTCAACGCTCG
 TGGCTCAGGGCTGGGTGGCAGAGGAGGCCCTTCACTTATATCTGTGTTGTTATCCAGGGCTCC
 AGACTTCTCTCTGCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
 AGCCCACTTGGCTTCTGTGCCCTCCTGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
 CAGAACACAGGGCCTGGCCAGTGGATTTATGGTGATCATTAATAAGAAAAATGCAACCAA
 AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSQVDICINNAGLARPDTLISGSTS~~SGWKDMFN~~VNVIALSICTR
EAYQSMKERNVDDGHIININSMGHRVLP~~LSVTHFY~~SATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYE~~QMKCLK~~PE~~DVAE~~VIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCTCCTGCTGCTGCTGCTCTTCTCTCAGCCTCCTGCCGGTGGCCTACACCAT
 CATGTCCCTCCACCCCTCCTTTGACTGCGGGCCGTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
 ACCTCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATCCAGTTCTGGTTTCATGC
 CAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCATGCCATTTAAGAGGGTTTTCTGCCA
 GGATGGAATGTTAGGTCGTTCTGTGCTGCGCTGTTTCATTTTCAGTAGCCACCAGCCACCTGTGG
 CCGTTGAGTGCTTGAAATGAGGAACCTGAGAAAATTAATTTCTCATGTATTTTCTCATTATTATTA
 TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
 TATATAATGATCAAAATCAGGGTAACCTGGGATATCCATCACAATCAAAATTTATTTTTATCTTT
 TTAGACAGAGTCTCACTCTGTCAACCAGGCTGGAGTGCACTGCTGCCATCTCAGCTTACTGCAAC
 CTCTGCCCTGCCAGGTTCAAGCATTTCTCATGCCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
 GCACCACAATGCCAACTAATTTTGTATTTTGTAGAGACGGGGTTTGCCATGTTGCCCAGG
 CTGGCCTTGAACCTCTGGCCTCAAAATCCACTTGCCTCGGCCCTCCAAAGTGTATGATTACA
 GGCGTGAGCCACCGTGCCCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAACCTTGAAATAT
 ACAATGAATTATTGTTAACTGTCACTCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
 ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCACTCCCACCTCTCTATCCTTCCC
 AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTAGCTCCACATGTG
 AGTAAGAAAATGCAATATTTGTCTTCTGTGCCCTGGCTTATTTCACTTAACATAATGACTTCTGT
 TTCCATCCATGTTGCTGCAAAATGACAGGATTTGTTCTTAATTTCAATTAATAATACCAACATG
 GCAAAAA

[illegible]

Important features of the protein:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCTCATCGGGACTAAAATTGGGCTGTTCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCTGCGATCGGGTTTCATTACTGTAATGATCGCTTCTGACATCCATTCACACG
GAATACCCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAATAAATATGCTGGGATCTCTTCAGAT
TTGAAAACCTTGTCTGAAAGTAGAAAGAAATACCTATACCAACACAGTTTAGATGAATTTCTACCAACCT
CCCAAAGTATCTAAAAGAGTTACATTTGCAAGAAAAATACATAAGGACTATCACTTATGATTCACCTTCAA
AATTCCTATCTGGAGAAATTACATTTAGATGACAACCTCTGTCTCTGCACTTAGCATAGAAGGGAGCA
TTCGAGACAGCACTATCTCCGACTGCITTTCTGTCCGCTAATCACCTTAGCACAAATTCCTGGGGTT
GCCAGGACTATAGAAGAATAAGCTTGGATGATAATCGCATATCCACTATTTTATCACCATCTCTTCAAG
GTCTCACTAGCTAAAACGCCTGGTTCTAGATGGAACCTGTGGAACAATCATGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAAATTTGACAGAGCTGTCCCTGCTGGGAAITTCCTGACTGCTGCACAGTAACCT
TCCAGGCACAACCTGAGGAAGCTTTATCTTCAAGTAACCACTCAATCGGGTCCCCCAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATTAATACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACCAACCTGATTTCTCGCAACAATCCCTGGTATTCGGGGTCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACCTGCTGGGGCTCATGTGCCAAGCCCCAGAAA
AGSTTCGTGGGATGCTATTAAAGATCTCAATGCAAGAACTTTTGATTTAGGACAGTGGGATTTGAAGC
ACCAITCAGATAACCACTGCAATACCCCAACACAGTGTATCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAACACGCCAGATATTAAAGAACCCAGCTCACTAAGGATCAACAAACCAAGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGCACCTCTGATACCAATCATATCTCTTGGAACTTGCTCACTATG
ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGCAITTTGGATCTAATACAGAAACAATTTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCCCTATAAAGTATGCACTGGTTC
CCATGGAAACCAAGCAACCTCTACCTATTTGATGAACTCCTGTTTGTATTTGAGACTGAACCTGCACCCCTC
CGAATGTACAACCTCACACCACTCTCAATCGAGAGCAAGAAAGAACCTTACAAAACCCCAATTTACC
TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCAITTGCCCTTCTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTTCAAGGAAGTGTGCAATAGCAAAAGGAGGAGGAAGAAGGATGACTAT
GCAAGCTGGCACTAAGAAAGCAACTCTATCTGGAAATCAGGGAATCTTTTTCAGATGTTACCAAT
AAGCAATGAACCACTCTCGAAGGAGGAGTTGTATACACACCAATATTTCTCCTAATGGAATGAATCTGT
ACAAAACAATCAGTGAAGGAGTGTATACCGAAGCTACAGACAGTGGTATTCACAGACTCAGATCAC
TCACACTCAATGATGCTGAAGGACTCACACAGACTGTGTITTTGGGTTTTTAAACCTAAGGAGGTGATG
GT

FIGURE 132

MISAASWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRLFITS IPTGIPEDATTLYL
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFFPTNLPKYVKELHLQENNIRITITYDSLSKIPYL
 EELHLDNSVSAVSIEBEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVNSLTAAVNLPGTNLRKLYLQDNHIN
 RVPPNAFSYLRQLYRLDMSNNNLSNLPGQIFDDLDNITQLILRNNPWCYCGCKMKWVRDWLQSLPV
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDGSIVSTIQITTAIPNTVYPAQGWPAFVTKQPD
 IKNEPLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
 IVTGERSEYLVTALEPDSPIKVCMPMETSNLYLFDETPVCIEETETAPLRMYNPTTTLNREQEKE
 PYKNPNLPLAAIIGGAVALVTIALALVWCYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
 ILEIRETSFQMLPISNEPISKEEFVIHTIFFPNGMNLKYNHSESSNRSYRDSGIPDSDSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGCA
GCCCTGCCGGGCCACTTGTCTTCA**ATG**CTCTGCGAGGGGAGGTGGGAAGSAGGTGGGAGGAGGGCG
TGCGAGGCGAGTCTGGGCTTGGCCAGAGCTCAGGGTCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCCTGGAGAGAGCCACCGTCTTCTTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCTGC
TGCGAGCCGTGAGCCTGCCGTGGGGTCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTCCAGCTGACCTCCAGCCCGG
GTTTGGAAAGCTCCACATGCTTGGATCCACACTGATGCTCCTTGGTGATACCCAGTTCGGGC
CCCAGGACTCATTTCTAGAGGAGAGAAGTGACGTGTGCTGGTGCGAGCTGCTGGGAACCGGACG
GACAGCAGCGAGCCCTGGGCCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCAACCACTGCTCTTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCARCCTCTCTCTCGCCAAACATGATGGACTTGAACCG
AGAGCTGAGGCCATCGGATACGCTACCTACCCGGGACATCTTCATGAAAAACATCATGTTCTG
TGGAATGGGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCAATTTTTCGAGGAGAGTGAAGAGCGAGAAAAACAAATTCAGAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCTGGGTTCAAGC
AATTCTCTTCCTCATCTCTCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACCTGGCTAAT
TTTTATATTTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAATCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTA**ATG**GTGTGAGCCACCGTGTCTG
GCTGGAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTG
ATGGGGGCTCTCTCCCTAGATGGCTGCTCTCCCAACAACAGCAGCAGCAGTGGCAGCCTGG
GTGGCTTCTATACATCTGGCAGATACCCGCCAGCAACAGAGAGCCACACCATCCACACCG
CCACCACAAGCAGCGCTGAGACGGACGGTTCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCTCATCCCTTAGATCTGGAGGGCAGGATCACATCTGGGAAGAAGGCATCTGGAG
ATAAGCAAGGCCACCCGACACCCAACTTTGGAAGCCTGAGTAGCGAGGCCAGGTAGGTGGG
GGCCGGGAGGACCCAGGTTGAACGGATGAATAAAGTTCACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPMASLGLLLLLLTALPPLWSSSLPGLD
TAEKATIAIDLILSALERATVFLEQRLPEINLDGMVGVRLVEEQIKSVREKWAQEPLLPQLSLRV
GMLGEKLEAAIQRSILHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHIDASLVYPTFGPQDSFSEE
RSDVCLVQLLTGTDSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGFLQSQSD
YINLFCANMDLNRRAEAGYAYPTRDIFMENIMFCGMGGFSDFYKLRLWLEAILSWQKQEGECFG
EPDAEDEELSKAIQYQQHFSRRVRKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGCTTCTTTCCCGTCTGCTGCTG
CTGCTGCTATCGGGGGATGTCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGCGTGCAGTTGTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCGAGTGCTGGTAGACGGAGAAGACACGTGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAGGGAATCGTGGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCAGAGCTGGCATTGGCAC
AAACACGGCACACTGGGTGGCATCCAAGTCTTGGAAAAACCGTGTGAAGCACTACTATAAACTT
GAGTCATCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACCTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAAATAAATTAT
ATGAACTACTATACATTATGTATATTAAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLALLLLSGDVQSSEVPGAAAEQSGSGVIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEHVGFLKTDGSEFVVHDIPSGSYVVEVVSAYRFDPPVRVDITSGKGMARYVNYIKTSE
VVRLPYPLQMKSSGPPSPYFIKRESWGWTDFLMNPMVMVMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSGSSKTKSGAGAKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 104-107

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCGAGTCCCGTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCGCCAAACAA
GTTTGTGACATTTCCCTGAAATGTCATTCTCTATCTATTCTAGTCAAGTGCCTGTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTTACGGGAGGCTTGGCAGT
TTTTCTTACTCTCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCTTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTCTGAGATACGGGCGAGTG
TGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTCTCGGCCATTTGCTAAGACTCTATCTGGACAGGATTTTAA
AACTACCAGACCCCTGACCATTATACTCTCGGAAGATCAGCAGCCTCGCCAAATCTCTTCTTA
CCATCAAGAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAATACAGCCAGATTCTGAGTCACTTTGAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGCG
TTTGGGGAACTAGACATTCTTCTGCAATGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTAICTTATTTATGCAATTACTTG
CTTCCTTGCAATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTGTATAGATTTT
TGTAATATCTTCTGCTATTGGATATATTTATAGTTAATATATTTATTTATTTTGTGCTATTTA
ATGTATTTATTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAACCTTGTAATTTCTAGAAGAGTGG
CTAGGGGGTTATTCATTTGTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATGAACCAATGACTACTTAGGATGGTGTGGAATAAGTTTGTGATGGGAATTCAC
ATCTACCTTACAATTACTGACCATCCCGAGTAGACTCCCCAGTCCCATATTTGTATCTCCAG
CCAGGAATCCTACACGCCACGATGTATTTCTACAAATAAAGTTTCTTGTGCATACCAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPVPGLTWCWALTAEPGWGQNGATTTCATNSHSDSELRFEIF
SSREAWQFFLLWSPDFRPMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLAKIS
SLANSFLTIIKDLRLSHAHTCHCGEAMKKYSQILSHFEKLEPQAQAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCGGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGTGGGGTCGGTTCGCCATCCAGCC
 TAGCGCTGTCCACAGTGGCGCTGGGGCTCCGGGACTTTCCGTACTGTTGCTAGCGATCGAGGTGC
 TAGGGATCGCGGTCTTCTCTCCGGGATTCTCCCGGCTCCCGTTCGTTCCTCTGCCAGAGCGGAA
 CAGGAGCGGAGCCCCAGCGGCCGAACCTCCGGCTGGAGCCAGTTCCTAAGTGCACAGCGCTGCC
 ACCACCTCTCTTACGTAAAGTTGTTATGTTCTGATAGATGCCCTTGAGAGATGATTTGTGTTTG
 GGTCAAGGGTGTGAAATTTATGCCCTACACAACTTACCTTTGTGAAAAAGGAGCATCTCACAGT
 TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCTCGAATCAAGGCATTGATGACGGGGAG
 CCTTCTCGGCTTTGTGAGCGTCATCAGGAACCTCAATTCTCTGCAGTGTGGAAGACAGTGTGA
 TAAGCAAGCAAAAGCAGCTGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
 TTCCCAAAGCATTTTGTGGAATATGATGGAACAACTCATTTCGTGTACAGATTACACAGAGGT
 GGATAATAATGTACAGAGGCATTGTGATAAGTATTAAGAGAGGAGATTGGGACATATTATATCC
 TCCACTACCTGGGGCTGGACCACATTGGCCACATTCAGGGCCCAACAGCCCCCTGATTGGGCAG
 AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
 GACGCTTTACCCAAATTTGCTGGTCTTTGTGGTGACCATGGCATGCTGAAACAGGAAGTCAG
 GGGCTCTCCACCAGGAGGTGAATACACCTCTGATTTTAAATAGTTCGCTTTGAAAGGAAA
 CCCGGTGATATCCGACATCCAAAGCAGTCCAAATAGACGGATGGCTGCGACACTGGCGATAGC
 ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCCTCTATTCCAGTTGTGGAAGGAAGAC
 CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCACTTAGTAACTGTTGCAAGAG
 AATGTGCCGTATATGAAAAAGATCCTGGGTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
 GAATGGATCAGACTGTACTTGGAGGAAAGCATTGAGAAGTCTATTCAACCTGGGCTCAAGG
 TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCCTGAGTGACAAAGTGAGCCAG
 TTCTCAGCTGCTCTGCTCAGCGTCCACAGGCACCTGCACAGAAAGGCTGAGCTGGAAGTCCCA
 CTGTCACTCTCTGGGTTTCTCTGCTCTTTTATTGGTGATCCTGGTTCTTTCGGCGGTTACAGT
 CATTTGTGTGCACTCAGCTGAAAGTTCTGCTACTTCTGTGGCTCTCTGTGGCTGGCGGAGGCT
 GCTTTCTGTTTACCACTCTGTTTGAACACCTGGTGTGTGCAAGTGTGGCAGTGCCTGGAC
 AGGGGGCTCAGGGAAGGACGTGGAGCAGCTTATCCAGGCTCTGGGTGCTCCGACACAGGTG
 TTCACATCTGTGCTGCAGGTGAGTGCCTCAGTTCTTGGAAAGCTAGTTCCTGCGACTGTTAC
 CAAGTGATTTGAAGAGCTGGCGGTACAGAGGAACAAAGCCCCAGCTGAGGGGGTGTGTGAA
 TCGGACAGCTCCACGAGAGGTGTGGAGCTGCAGCTGAGGGAGAGAGACAATCGGCTTGGA
 CACTCAGAGGGTCAAAGGAGACTTGGTCGACCACTCATCTGCCACCCCAAGATGCATCT
 GCTCATCAGGTCAGATTCTTTCOAAGGCGGACGTTTCTGTGTGAATTCTTAGTCTTTGGCC
 TCGGACACCTTCATTGTTAGCTGGGGAGTGGTGTGAGGCAAGTGAAGAGGCGGATGTCAC
 ACTCAGATCCACAGAGCCAGGATCAAGGACCCACTGCATGTCAGCAGGACTGTGGGCCCCC
 ACCCCAGCCCTGCACAGCCCTCATCCCCCTTTGGCTTGAGCCGTCAGAGGCCCTGTGCTAGTGT
 CTGACCGAGACACTCAGCTTTGTATCAGGGCAGAGGCTCTCTCGAGCCAGGATGATCTGTG
 CCACGCTTGACCTCGGCCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
 CTGCACACAGTATGTAGTTACCAAAAGATAAACGGCAATTAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSGVKFMPYTTYLVKKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHIGHISGPNPLIGQKISEMDSVLMKIHTSLQSKERETPLP
NLLVLGCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGSTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACTACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACCGCGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCGGATGTTTCATCTGCAOCTCCTGCAATTGTAATGAGCCT
GTTGGSGTGACAGATAAATTGAGAACAGGAACACATTGAATTTTCATTTC AACAGTTTGCAA
AGCTGAAATGAGCCCACTGAGGTGAGCGATAGGAAACTGCCCATTTGAACGCCTTCCTCGCTA
ATTTGAACATAATTGTATAAAAAACACCAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCOHLKTHTFRVKNLNPKKPSIHDQDHKVLVLDSGNLIAPDKNYIRPEI
FFALASSLSASAEEKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCTTCTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTRAACCTCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGGCTGTCTGTCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCAATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACTGCTTCTTACCATTGACACAGGCCCTGGATGACTCAGITCAGC
CTCTGAACAAGACTGTCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCTGGACATCAGCACTACTCTCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTGTTAGAAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAA
AAA

FIGURE 144

MLGLFWKGGSLWALLLLLSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHIILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLINKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAAACCGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAAACCGCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCGATGGCATCCA
CTACACAGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACACGGCTGCATCAATGCCA
CCGAGCGCGCAACGAGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTCGAGAGTTTGGTTGAGAGGGGCGCAGGACT
TCGGGTACCATGCAACAGCCAGTGTCTCTGCTTCTGGCTTTGATCTGGCTCATGGTGAAT
AAGCTTGCCAGGAGGCTGSCAGTACAGAGCGCAGCAGCGAGCAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGCGATGCATCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTTAGCAGAGCGCTGTGGCACACTAGATTAGTAGTAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTACAGTACTTCCCAACACTCTTAGAGGTAG
GTGTATTCCGTTTTACAGATAAGGAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTTGTCTAACCCAGGTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAATGTGAAGCTCCATGTTTAGAAATAAATGAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHKRIKWNRKALPSTAQITEAQAENRFGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAAHQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLAVTMHQPVLLCLLALIWMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGGTGACCACCATGCTGGTGCAGAGTCTCATTT
CTCTGATGATTATAGTACTCAAAGAAACTCAATGTTTCAGAAGTCTCTCTCTCTCTGGCTCCTCTCT
CTGTGCTTTCTCCCTCTCTCTCTTTATTATAGTACGTCACTCAGAGTCATGCAGCAGTCGG
AAATCTTTGCATTTTGTCTGTCTAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTGAAATT
CAACTTTTCAGATTACAGGGGGTACATGTGAAGGTTTGTTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFLQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 149

GTCTCCGGCTCACAGGAACCTCAGCACCCACAGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
 TCCCGCGCGCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCGCCCACTGCC
 CGGCCCTCTCCCGCCCCACACCACCCCTCTGCGCTCTCTGCTTTTACTCCTCCTTTTCATTGATCATA
 ACAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAGCCGAGCGTGAAGAATGCGGGTT
 CCTCGGACCGGCACCTGGATTCTGGTGTAGTGTCTCCGATTCAAGCTTTCCCAACCTGGAGGAA
 GCCAAGACAAATCTCTACATAATAGAGAATTAAATGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
 GCAGAAGAAGACAAGATTAACCAACATATCCTCCAGAAAAACAAGCCAGGTGAGAGCAACTATTCTTT
 TGGTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAAGGAAAAAATTGAGAAAAAGACRACTTA
 TAAGAAGCTCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG
 ATCGATGATTATGACTCTACTAAGACTGGATTGGATCATAAATTTCAAGATGATCCAGATGGCTTCA
 TCAACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
 AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCCTTATCAGAAAAAG
 CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAAATTAATCTCAAGGAAGCCCAAA
 TTATGAGGAGGATCCCAATAAGCCCCAAGCTGGACTGAGAATCAGGCTGAAAAAATACCAGAGAAAG
 TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGAGAAAAACGATGAACAGTATCTAACACA
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA
 ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGAGCAAAAGAGAAAGAAA
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
 CCAGAAGAAGTGTTTCTACCTTGAAAACCTGGATGAATGATTGCTCTTCAGACCAAAAAACAAGCT
 AGAAAAAATGCTACTGACATAATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTATGAAGAAA
 CAGACAGTACCAAGGAAGGAAGCAGCTAAGATCGAAAAAGAAATAGGAAGCTTGAAGGATTCCACAAAA
 GATGATAAATCCCAACCCAGGAGGAAGACAGATGAACCAAGGAAAAACAAGAAGCCTATTGGAAGC
 CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGAAAAATAAGAAGATTATGACCTTT
 CAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAGGAA
 GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAATATGGCAAAAGATCCAGGAGCTTTTCAA
 CTGTTTCAGAAAAACATAATATAGCTTAAACACTTCTAATCTGTGATTAAAAATTTTGGACCAAGS
 GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
 GTAAAACTATCTGAAGTAAAGTTGTATGTAAAGCTGAAAAAAGAAAAAAGAAAAA

FIGURE 150

MGFLGTGTWILVVLPIQAFKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERSIRSSPLDNKLNVEDVDSTQNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDFNKPTSWTENQAGKIPKVTFMAIQDGLANGENDTVSNTLTLTNGLE
RRTKTYSEDNFEELOQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISP EEGV
SYLENLDEMIALQTKNLEKNATDNISKLFPPAPSEKSHEETDSTKEEAAKMEKEYGSLKDKSTKDD
NSNPGGKTDEPKGRTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGAACATTCTGAGGGGAGTCTACACCTGTGGAGCTCAA
GATGGTCTGAGTGGGGGCGTGTGCTTCGAATGAAGACTCGGCATTGAAGTGCTTTATCTGC
ATATAACACAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGTGGCTGGATGCCAGCCTGTCCCCCTCATCTCTGGGTGTCCAGGGTGGAAAG
CCAGTGCCCTGCATGTGGGTGGGGCAGGAGCCGACTCAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCNAGGAATCCNAGAGCTTCACTTCTACCGGCGGACATGGGGCTCACC
TCCAGCTTCGAGTGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTGAGACTCACCAGCTTCCCAGAGATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCACTGCCGCCCAGAACTCCCTGGGAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGCCCATGGCGGACAACTCACTCTCTGCTCTCAGGACCCCCACGCTCTGACTTAG
TGGGCACTGACCACTTTGCTTCTTGTGTCCAGTTTGGATAAATCTGAGATTGGAGCTCAGT
CCACGGTCTTCCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTAAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAATTCTGTGGGGGTGGGTGGGGAGTGGTGGGAATCATCTCTGCT
TAATGGTAAGTGAACAGTGTACCTGAGGCCGCGAGGCCAACCCATCCCAGTTGAGCCTTATA
GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCAACCACTGTGAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCGAGCCCCACCCCTTCCCTTTAATCTGCACTGTCTATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTGTGTTGTGGGCATGAGGAGGTGGTATGTCAGAA
GAATGGCTGAGCTCAGAGATATAAGATAGTAGGGTATGCTGATCTCTTTTAAAAACCCAA
GATACAATCAAAATCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTATTATTAATTAAGATACCTATTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTTAGCAATGTGACGGGTGGTGGCAGTAT
AGGTGATTTTCTTTTAACTCTGTTAATTTATCTGTAATTTCTCAATTTTCTACAATGAAGATGA
ATTCTCTGTATATAAATAAGAAAAAATAATCTTGAGGTAAAGCAGAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCAGAGTAATTCAAATGAATCGAGCTCTGCTGCTTGGTGG
TTGTAGTAGTATCAGGAACAGATCTCAGCAAAAGCACTGAGGAGGAGGCTGTGCTGAGTTTGT
TTGGCTGGAACTCTGGGTAAAGAACTTAAAGAACAAATCATCTGTAATTTCTTCTAGAG
GATCACAGCCCTGGGATTCAGGAGTGGATCCAGTCTCTAAGAGGCTGCTGTACTGGTGA
ATTGTGCCCCCTCAAATTCACATCTTCTTGGAACTCTAGTCTGTGAGTTTATTGGAGATAAG
GTCTCTGCAGATGTAGTGTAGTTAAGCAAGGTCTAGTGTGATGAGGTAGACCTAAATCAATAT
GACTGGTTTCTTGTATGAAAGGAGAGGACACAGAGACAGAGGAGAGCGGGGAAGACTATGTA
AAGATGAGGCAGAGTCCGAGTTTTCAGCCCAAGCTAAGAAACCAAGGATGTGGCAACC
ATCAGAGCTTGGAGAGGCCAAGNAGATCTCTCCCTAGAGGCTTTAGAGGATAACGGCTCTG
CTGAACCTTAACTCAGACTTCCAGCTCTCTGAACGAAAGAAATAATTTGGGCTGTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTTAGGAACTAATACAGCTGTCAAAATGATCCCGT
CTCCTCGTGTACATCTGTGTGTGCCCTCCACAAATGTACAAAGTTGCTTTGTGACCAA
TAGATATTGGCAAGATGATGGCATGCCACTTCCAAGATTAGGTATAAAGACACTCGAGCTTC
TACTTGAGCCCTCTCTCTGCCCCCACCACGCCCAACTCTACTTGGCTCAGCTCGCTCGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAGAGACTTACGTGGTAAAAAATGAAGTCTCT
GCCCAGGCCCACTATTGTAACCTAGAAGCAGAGCTCTGAGATAATCGATGTTTGGTTT
AAGTGTCTCAGTTTGGTCTAATCTGTTATGAGCAATAGATAAATAATATGCAGAGAAAGG

FIGURE 152

MVLGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEETISVVPNRWLDASLSPVILGVQGGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFIFYRRDMGLTSSFESAAYPCWFLCTVPEADQP
VRLTQLPENGWGNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGCTCGCAATGCGCGC
CCTGCAGAAATCTGTGAGCTCTTCCCTTATGGGGACCCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCACAGCCCCATATACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTGCTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTC
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGCTGGATTTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCTGCTAGAAAAATA
CAATTAGATGCCCAAAGCGATTTTTTAAACCAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTATTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAACCTGGTGTC
TATACACAGAAAACAAATTTATTTTTAAATAATGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLDKTVRKLGESGEYKAIGELDLLFMSLRNACT

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGC TGAARATAAAATCAGGACTCCTAACCTGCCTCCAGTCAGCCTGCTTCCACGAGGCCCTGT
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCGAGCATGTACCAGGTCAGTCAGAGGGC
TGCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGGAGGCCAAGC
TGCCAGGT TTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCCTCATTAGCCTTTTCTACAGGTGGTTCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCTCCCTAGAGCCTGCTAG
GCCCCAACGCCACCCAGAGTCCGTGAGGGCCAGTGAAGATGGACCCCTCACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCGT
TGCC TGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA
GCTGCTCTACCAACCACTGCTTCTACAGCGGCCATGCCATGGCGAGAAGGGCACCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAACTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGTGGAGCAG
CAGGATCCCGGACAGGATGGGGGGCTTTGGGGAACCTGCACTCTGACATTTTGAAGAG
CAGTGTCTGCTTAGGGCCCGGAAGTGGTGTCTGTCTTTCTCTCAGGAAAGGTTTTCAAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTCTCTCTTTTCCCATCCCTGCTACCCCTG
GCCAGCAGCAGGCACTTTCTAGATATTTCCCTTGTGAGGAGAAAGAGCCCTGGTTTTATT
TTGTTTGTACTCACTCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTTATTTAAAAATGAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCCPKQDTSEELLRWSTVPVFPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVCRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGAAGCCTGGCGCGCTGTGCAGGAGCGCGGTACCCCGAGAGCC
 GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
 CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAC TAGTGTGCAACAGGGGACTATTCA
 ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCGCTTGTGAAGGCCACCAA
 GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTTACAGCTGTGTGAGGTGCAATTACACAGAGG
 CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTTCTTACATCGGCTTCCCTGTGA
 GAGCTGAACACAGTCTATTTCAATGGGGCCATAATATTCTTAATGCAATATGAATGAAGATGG
 CCCTTCCATGCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
 AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCAGTCTTGTGAAGAATGAGGAGACA
 GTAGAGTGAACCTCACAAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAAACACAGCAC
 TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
 TTCCAGTGACTGGGATAGTGAAGTGTCTACGGTGCAGCTGACTCCATATTTTCTACTTTGTGGC
 AGCGACTGCATCCGACATPAAGGAACAGTGTGTCTCTGCCCAACAAACAGGCGTCCCTTTCCCTCT
 GGATPAACAAACAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCCTCTGCTCTGTCTGTGTGGTGC
 CATGGGTCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAGGATCAAGAGACTTCCTTT
 TCTACCACCACACTACTGCCCCCAATTAAGGTTCTTGTGTTTACCCATCTGAAATATGTTTCCA
 TCACACAATTTGTACTTCACTGAATTTCTTCAAACCATTTGCAGAAGTGAGGTCATCCTTGAAA
 AGTGGCAGAAAAAGAAATAGCAGAGATGGGTCAGTGCACTGGCTTGCCTCAAAAGAGGCA
 GCAGACAAAGTCGTCTTCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
 GAGCGAGGGCAGTCCCACTGAGAATCTCAAGACCTCTTCCCTCTGCTTAACTTTTCTGCA
 GTGATCTAAGAAGCCAGATTATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
 AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
 CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCAACAGCTGCCACG
 ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGETGSPENMLQHDLPGLKDLRVEPVTTSVATGDYSILMNVSVV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQOTRPSGGKWTFSYIGPPVELNTVYFIGAHNIP
NANMNEDEGSPMSVNETSPGCLDHIMKYKKKCKVAGSLWDPNITACKNKEETVEVNFTTTPLGNRYMALIQH
STIIIGSQVFEPHQKKQTRASVVIPTVGDEGATVQLTPYPPTCGSDCIRHKGTVVLCPTQGVFPPLDNKK
SKPGGWLPLLLSLLVATWVLVAGIYIMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICPHHTICYTFEFL
QNHCRSEVILEKWKQKKIAEMGFVQNLATQKKAADKVVLLSNDVNSVCDGTCGRSEGGSPSENSQDLFELA
FNLFCSDLRSQIHLAKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVQVQSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

[illegible][illegible]

FIGURE 160

MTVKTLHGPA¹MPV²KYLLLS³ILGLAFLSEAAARKIPKVGH⁴TF⁵QKPE⁶SCPPV⁷PGGSMKLDIGI⁸INEN
Q⁹R¹⁰V¹¹S¹²M¹³R¹⁴N¹⁵IE¹⁶S¹⁷R¹⁸T¹⁹SPW²⁰NY²¹TV²²WD²³PN²⁴RY²⁵PSE²⁶V²⁷V²⁸QA²⁹Q³⁰R³¹N³²LG³³C³⁴INA³⁵Q³⁶G³⁷KE³⁸DIS³⁹M⁴⁰NS⁴¹V⁴²PI⁴³Q⁴⁴ET⁴⁵L⁴⁶V⁴⁷
RRKH⁴⁸QG⁴⁹CS⁵⁰VS⁵¹F⁵²Q⁵³LE⁵⁴R⁵⁵V⁵⁶LV⁵⁷TG⁵⁸CT⁵⁹CV⁶⁰TP⁶¹VI⁶²HH⁶³V⁶⁴Q⁶⁵

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

FIGURE 162

MPVFWFLSLALGRSPVVLSERLVGFQDATHCSFGLSCRLWSDSDILCLPGDIVPAGFVLAPTHLQTELV
LRQKQETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSPQAYPTARCULLEV
QVFAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTPQPRYEKELNHTQQLPALFWINVSADGDNVHLVLMVS
EEQHFGLSLYWNQVGPPKFRWHKNLGTGQIITLNHTDLVPCLCIQVWFLEPDSVRNICFPREDPRAHQN
LWQAARLRLLTLQSWLLDAPCSLPAAALCWRAPGGDPQQLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQELWADSLGPKDDVLLLETRGPQDNRSLEALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGGPVANFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGGAGHPDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLLHPDAVPALFRTVPVFT
LPSQLPDFLQALQPPRPRSGRLQERAEQVSRALQFALDSYFHPFGTPAPGRGVGFGAGGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
GCTCAGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCAGTGAATTCAGTCCAGCAACTTTGA
AAACATCTTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGTGGCAAGAAGGGCTGTACGGGATCACCCGGAAGTCTCTGC
AACTGACGGTGGAGACGGCAACTCAGCGGAGCTCTACTATGCCAGGGTCACCCGT
GTCACTGGGGAGGCCGTTGAGCCACCAAGATGACTGACAGGTTGAGCTCTTGAGCACACTAC
CCTCAGGCCACTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGTTGTCATCCTA
CCCCACGCCAATCCGTGAGGGGATGGCCACGGCTTAACCTGGAGACATCTTCCATGACCTG
TTCTACCATCTAGAGCTCAGGTCACCGCACCTACCAATGCACCTTGGAGGGAAGCAGAGAGA
ATATGAGTTCTTCGGCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGGTGTCCCA
CCTGGGCCAAGGAGGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCTCTCCGGAGCTTCTTGTCTCCATGGGCTTCTCTGTCGCAGTACTCTGCTACCTGAG
CTACAGATATGTACCAAGCGGCTGCACCTCCCACTCTCTGACGTCACGGAGTCTTGACTT
TCAGCGCTGCGCTTTCATCCAGGAGCAGCTCTGATCCCTGTCTTTGACCTCAGCGGCCCGCAGC
AGTCTGGCCAGCCTGTGACGACTCCAGATCAGGGTGTCTGGACCGAGGAGCGCCGAGGAGC
TCCACAGCGGCATAGCTGTCCGAGATCAGCTACTTAGGGCAGCCAGACATCTCCATCTCCAGC
CTCCAACTGCCACTCTCCAGATCTCTCCCACTGTCTTATGCCCAAACGCTGCCCTGAG
TCCGGGCCCATCTCTATGCACTCAGGTGACCCCGAAGCTCAATCCCATTCTACGCCCCACA
GGCCATCTCTAAGGTCAGCCTTCTCTATGCCCTCAAGCCACTCCGACAGCTGGCTCCCT
CCTATGGGGTATGCATGGAAGTTCTTGGCAAAGACTCCCGACTGGGACATTTCTAGTCTTAA
CACCTTAGGCTTAAAGTCTGAGTTTCAAGAAAGCCACAGCTGGAAGTGCATTTAGGTGGCT
TTCTTCGAGGAGGTGACCTCTTGGCTATGGAAGAAATCCCAAGAACAAATCATTTAGCCAGC
CCCTGGGATTTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGAGCA
CCACAGTACCTTAAAGGGCCAGCTCCCGCTCTCTCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTTGCACTCTCTTCGGTCCATGTTCCCTCGGACCAAGGTCCAAGTCCCTGGGGCT
TGCTGGAGTCCCTTGTGTGCCAAGGATGAAGCCAGAGGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAACTGGATTCTCTTTTCAAGGCTCGGCTGACTGTGCTGAGTGGGAGTCTCTG
AGGGGAATGGGAAAGCTTGGTGCTTCTCCTGTCTCTACCCAGTGTACATCTTGGCTGTCA
ATCCCATGCTGCCCATGCCACACTCTGCGATCTGGCTCAGACGGGTGCCCTGAGAGGAAGC
AGAGGGATGGCATGAGGGGCCCTGCCATGGGTGCGCTCTCACCGGAACAAGCAGCATGATA
AGGACTCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGAACAGCGCTGCTCTGCTGAGCCCTG
CAAGGCAGAAATGACAGTGAAGGAGGAAATGACGGGAACCTCCGAGTCCAGAGCCCACTCT
CTAACCACTGATTCAAAGTGCTCAGGGAATTTGCTCTCTTGCCTCATTTCTGGCCAGTTTC
ACATCTAGCTGACAGAGCATGAGGCCCTGCTCTCTCTCTGCTCATTTCTCAAGGTGGGAAGAGA
GCCTGGAAGAAAGACCGGCTGGAAAGAACAGAGGAGGCTGGGACAGACCAAGAACACCTGC
ACTTCTGCCAAGGCCAGGGCAGCAGGACGGCAGGACTTAGGGAGGGGTGGGCTGCGATCTA
TTCACAGCCAGGGAACCTGCTGACGTTGACGATTTGACGTTCTTCTCTGATAGAACAAAGC
GAAATGCAGGTCCACAGGAGGGAGACACAGCCCTTTTCTGCGAGCAGAGTTTTCAGACTT
ATCTCAGATAGGGGTTTGAAGGAAGGTGAGGCTGTGGCCCTGGACGGGTACAAATACACAC
TGTACTGATGTCAACTTTGCAAGCTCTGCTTGGGTTCAGCCCATCTGGGCTCAAAATTCAGC
CTACACCTCACAAGCTGTGTGACTTCAACAAATGAATCAGTCCCAAGACCTGGGTTTCTCT
ACTCTAATGTGGGATCATACACCTACCTCATGAGTGTGGTGAAGATGAATGAAGTCAATG
TCTTTAAGTGCTTAATAGTGCCTGGTACATGGGACGCGCAATAAAGCGTAGCTATTAAAAA
AAAAAA

FIGURE 164

MRTLTLITVGSAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDVTVYSIEYKTYGERDW
 VAKKGCORITRKS CNLTVETGNLT ELYYARVTVASAGGRSATKMTDRFSSLQHTTLTKPPDVTCIS
 KVRSIQMI VHP TPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLLGGKQREYEFFGLTPDT
 EFLGTIMICVPTWAKESAPYMC RVKTL PDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
 PNSLNVQRVLTFQPLRFIQEHVLI PVFDLSGFPSSLAQPFVQYSQIRVSGPREPAGAPQRHSLSEIT
 YLGQPDISILQPSNVPPPILSPLSYAPNAPEVGPSPSYAPQVTPEAQFFFYAPQAISKVQPSY
 APQATPDSWPPSYGVCMESGKDSPTGTLSSPKHLRPFKGLQKEPPAGSCMLGGLSLQEVTSLAM
 EESQEAKSLHQPLGICTDRTSDPNVLHSGEETPQYLKQQLPLSSVQIEGHMMSLPLQPPSGPC
 SPSPDQGPSFWGLLES LVCPKDEAKSPAPETSDLEQPTELDLSLFRGLALTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

[illegible][illegible]

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDTFGPDCRFVN
FKKGDDVYVYKLAGGSLELWAGSV EHSFGYFPKDLIKVLHKYTEELHIPAETDFVC FE GGRD
DFNSYNVEELLGSLELSDVP EESKKAEEVSQHREKSPEESRGRELDPVPEAFRA DSEDEGEA
FSESTEGLQGQPSAQESHPHTSGPAANAQGVQSSLDTFEEILHDKLKVPGESERTGNSSPASVER
EKTDAYKVLKTEM SQRGSQCQVIHYSKGFWRHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCTCTAGGACATACACGGGACCCCTTAACCTC
AGTCCCCCAAACGCCACCCCTCGAAGTCTTGAACCTCCAGCCCCGCACATCCAGCGCGGCACAGG
CGCGGCAGGCGCGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCTGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTTCGCGGGCTCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCAATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCGCGCGCGTGGTCTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAAATGGCCTACTTCCATGAACT
GTCCAGCCGAGTGAGCTTTTCAAGGAGCACGCTGGCTTGTGAGAGTGAGGGAGAGTCCCTCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAAACATCTGG
TGCTTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCACAGTACCGAAACTGGTACACAG
ATGAACCTTCTGCGGAAGTGAAGTGTGTTGTGATGTATACCAACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACCAGTGAATGATGACAGGTGTAACATGAAGCACAATTTATTTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTAGAGAAAGCCCTTATCTTACAATCAACAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACTAGTCCAAACAGTCTACACTGTGGGATTTCAAAGAGTA
CCAGAAAAAGAAAGTGGCATGGAAGTATATAAATACTATTGACTGGTTCCAGAATTTTGTAATCT
GGATCTGTATAAGGAATGGCATCAGAACATAGCTTGAATGGCTTGAATCACAAGGATCTGC
AAGATGAAGTGAAGTCCCCCTTGAGGCAATATTAAGTAATTTTTATATGCTATTATTTCA
TTTAAAGAAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCCGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTCACCTTTTATAAGTTGTTATCTAGTCAATGTAA
TGTAATATGTTATGAAATTTACAGTGTGCAAAAGTATTTACCTTTGCATAAGTGTGTTGATAAAA
ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACAAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAAATTCATCTTTTTCAGCTTCTGCTTTGCTCAATGCTTAGGAAATCTCTTCAGA
AATAAGAAAGCTATTTCAATAGTGTGATATAAACCCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGCTCAATTTCAATTGTGCAAGACATGTGCTTATAATTTATTTTAGCTTAAATTTAAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAAG
TGACATACACAATATAATCATATGCTTTCACACGTTGCCTATATATGAGAAGCAGCTCTCTGA
GGGTTCTGAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGTTGTTGCGGGTTTGGG
ATTGACACTGGAGGCAGATAGTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTGATGTTGAGACACAGGTGAATAGTCACATATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTAGGAAGGAAGGAACTACGAAATCGTGTGAAATGGGTGG
AACCCTCAGTGAATCGCATATTTATTGATGAGGGTTTGCTTGAGATGAGAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT
CTGCTCTTATATGCTTCCCTGCTCCTTTTAAACCAAATAAAGAGTTCTTGTCTTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYKMAFYHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDWFGLWRNGDGGTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSkgRRTKTSNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217